



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY



**FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH
TECHNOLOGIÍ**

ÚSTAV BIOMEDICÍNSKÉHO INŽENÝRSTVÍ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION
DEPARTMENT OF BIOMEDICAL ENGINEERING

SHLUKOVÁ ANALÝZA SIGNÁLŮ EKG

CLUSTER ANALYSIS OF ECG SIGNALS

BAKALÁŘSKÁ PRÁCE

BACHELOR'S THESIS

AUTOR PRÁCE

AUTHOR

TOMÁŠ DVOŘÁČEK

VEDOUCÍ PRÁCE

SUPERVISOR

doc. Ing. JIŘÍ KOZUMPLÍK, CSc.

BRNO 2013



VYSOKÉ UČENÍ
TECHNICKÉ V BRNĚ

Fakulta elektrotechniky
a komunikačních technologií

Ústav biomedicínského inženýrství

Bakalářská práce

bakalářský studijní obor

Biomedicínská technika a bioinformatika

Student: Tomáš Dvořáček

ID: 133963

Ročník: 3

Akademický rok: 2012/2013

NÁZEV TÉMATU:

Shluková analýza signálů EKG

POKYNY PRO VYPRACOVÁNÍ:

1) Prostudujte problematiku shlukové analýzy vycházející ze vzájemných vzdáleností (či podobností) jednotlivých objektů. 2) Vyberte vhodné příznaky a navrhnete metodu pro automatickou shlukovou analýzu signálů EKG s cílem rozlišit normální cykly a různé typy abnormálních cyklů v signálech EKG. 3) Realizujte metodu shlukové analýzy signálů EKG v prostředí Matlab. K testování použijte vybrané reálné signály s výskytem abnormálních cyklů, např. z CSE databáze dostupné na UBMI. 4) Vypracujte studii shrnující výsledky práce.

DOPORUČENÁ LITERATURA:

[1] RANGAYYAN, R.M.: Biomedical Signal Analysis. A Case-Study Approach. IEEE Press, John Wiley & Sons, 2002, ISBN 0-471-20811-6

[2] ROMESBURG, H., CH.: Cluster Analysis for Researchers. Lulu Press, North Carolina, 2004.

Termín zadání: 11.2.2013

Termín odevzdání: 31.5.2013

Vedoucí práce: doc. Ing. Jiří Kozumplík, CSc.

Konzultanti bakalářské práce:

prof. Ing. Ivo Provazník, Ph.D.

Předseda oborové rady

UPOZORNĚNÍ:

Autor bakalářské práce nesmí při vytváření bakalářské práce porušit autorská práva třetích osob, zejména nesmí zasahovat nedovoleným způsobem do cizích autorských práv osobnostních a musí si být plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č. 40/2009 Sb.

Prohlášení

Prohlašuji, že svou bakalářskou práci na téma Shluková analýza signálů EKG jsem vypracoval samostatně pod vedením vedoucího bakalářské práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury na konci práce.

Jako autor uvedené bakalářské práce dále prohlašuji, že v souvislosti s vytvořením této práce jsem neporušil autorská práva třetích osob, zejména jsem nezasáhl nedovoleným způsobem do cizích autorských práv osobnostních a jsem si plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č. 40/2009Sb.

V Brně dne 28. května 2013

.....
podpis autora

Poděkování

Děkuji vedoucímu bakalářské práce doc. Ing. Jiřímu Kozumplíkovi, CSc. za veškerou jeho pomoc a cenné rady při zpracování mé bakalářské práce.

V Brně dne 28. května 2013

.....
podpis autora

Abstrakt

Tato bakalářská práce se zabývá problematikou automatické shlukové analýzy EKG signálu. V bakalářské práci je rozebrána teorie patologie srdečního rytmu a problematika shlukové analýzy. Na základě získaných znalostí je pak dále vytvořena sada příznaků, díky nimž je shluková analýza provedena v prostředí MATLAB a jejíž výsledky jsou zhodnoceny.

Klíčová slova

EKG, shluková analýza, QRS komplex, dendrogram, RR interval, metody shlukové analýzy, UPGMA, SLINK, CLINK, WARD, Euklidovská vzdálenost, Pearsonův korelační koeficient, vlny, MATLAB, plocha pod křivkou.

Abstract

This bachelor thesis is engaged in problematics of automatic cluster analysis for ECG signals. This bachelor thesis contains theory of pathology in cardiac rhythm and problematics of cluster analysis. Based on a gained knowledge, set of symptoms is made, which are used for cluster analysis using MATLAB interface and which results are evaluated.

Keywords

ECG, cluster analysis, QRS complex, dendrogram, RR interval, cluster analysis methods, UPGMA, SLINK, CLINK, WARD, Euklidean distance, Pearson's correlation coefficient, waves, MATLAB, area under the curve .

Bibliografická citace

DVOŘÁČEK, T. *Shluková analýza signálů EKG*. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2013. 64 s. Vedoucí bakalářské práce doc. Ing. Jiří Kozumplík, CSc..

Obsah

1.	ÚVOD	8
2.	EKG.....	9
2.1	Srdeční automacie – elektrická aktivita srdce.....	9
2.2	Snímání EKG	10
2.3	Vlny v EKG	12
2.4	Poruchy srdečního rytmu	13
2.4.1	Některé typy arytmií.....	13
3.	Shluková analýza.....	15
3.1	Úvod do shlukové analýzy	15
3.2	Postup u aglomerativních metod shlukové analýzy	16
3.2.1	Získání matice dat	16
3.2.2	Standardizace dat.....	16
3.2.3	Výpočet matice podobností	17
3.3	Provedení vlastní metody shlukové analýzy	18
3.4	Výpočet Pearsonova korelačního koeficientu	22
4.	Vybrané příznaky a jejich zdůvodnění	23
4.1	Tvarový faktor.....	23
4.2	Vzdálenostní faktor	24
4.3	R-R interval	24
4.4	Plocha pod křivkou QRS komplexu.....	24
5.	Obsah a obsluha programového řešení.....	25
6.	Zhodnocení výsledků práce.....	26
6.1	Použití atributu R-R intervalů a ploch QRS komplexů.....	28
6.1.1	Ukázky úspěšného shlukování	28
6.1.2	Ukázky neúspěšného shlukování	31
6.2	Použití atributu R-R intervalů a tvarového faktoru.....	36
6.2.1	Ukázky úspěšného shlukování	36
6.2.2	Ukázky neúspěšného shlukování.....	40

6.3 Poznámka k nastavení prahu řezu dendrogramem a ke speciálnímu způsobu rozměření některých signálů.	42
7. Závěr.....	46
8. Použitá literatura	48
Seznam obrázků	50
Seznam tabulek	51
Seznam zkratk, symbolů a příloh	53
Přílohy	54

1. ÚVOD

EKG jakožto elektrokardiogram je už několik desítek let nezbytnou terapeutickou i diagnostickou technikou, pomocí níž můžeme sledovat a posuzovat abnormality v srdečním rytmu a pomocí níž dokážeme diagnostikovat poruchy srdeční motoriky.

Vlny (P, Q, R, S, T) v EKG signálu zastupují jednotlivé části srdečního cyklu, a jejich tvar, vzdálenosti, výška a frekvence jsou důležitými ukazateli poruch tohoto cyklu.

Tato práce pojednává o shlukové analýze výše zmíněných cyklů v EKG a jejich rozdělení do skupin pomocí rychlých a jednoduchých metod shlukové analýzy. Proto se tato práce bude zabírat hlavně funkcí těchto vln v signálu EKG, jejich použitím pro detekci abnormalit v signálu a následným roztržiděním.

Aby byla analýza úspěšná, je nutné se dopodrobna zabírat funkcí jednotlivých vln v srdečním cyklu. K tomuto vysvětlení slouží podkapitola 2.3. Shluková analýza v tomto případě vyžaduje stanovit pro každý objekt (ve shlukové analýze EKG jsou objektem QRS, případně PQRST komplex) sérii příznaků, které jsou specifické pro jednotlivé abnormality v signálu, a pomocí těchto příznaků a aplikací metod shlukové analýzy je dále roztržidit do skupin na normální a abnormální cykly.

Práce tedy dále pojednává o výběru vhodných příznaků (atributů) pro shlukovou analýzu a jejich následné využití za pomoci programu, jež je vytvořen v prostředí MATLAB. Práce dále obsahuje výsledky jednotlivých analýz pro 21 signálů EKG z databáze CSE, obsahujících patologické srdeční cykly a odůvodnění příčin úspěchu, případně neúspěchu analýzy. Všechny výsledky jsou shrnuty v kapitole č. 6.

Celkové zhodnocení práce je nakonec shrnuto v závěru (kapitola č. 7) a analýzy, o jejichž výsledku nebylo pojednáno v kapitole č. 6, jsou graficky znázorněny v příloze této práce. V příloze jsou také zmíněny hodnoty Pearsonových korelačních koeficientů pro jednotlivé signály a pro jednotlivé kombinace atributů a metod, využitých k analýze těchto signálů (tab. 14). Přílohy dále obsahují blokové schéma programu, jenž je využit k provedení shlukové analýzy (obr. 47).

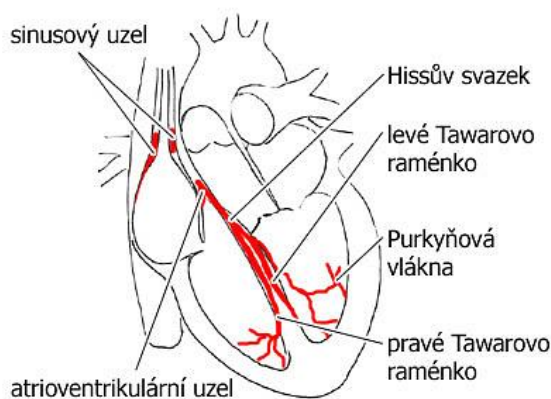
2. EKG

EKG neboli elektrokardiografie je důležitý diagnostický prostředek u získávání anamnézy a fyzikálního vyšetření. Jedná se tedy o základní vyšetření pro diagnostiku poruch rytmu a její použití je možno rozšířit i na diferenciální diagnostiku např. bolestí na hrudi.

2.1 Srdeční automacie – elektrická aktivita srdce

K pravidelnému střídání systoly a diastoly je zapotřebí, aby v tzv. převodním systému srdečním vznikala vzrušová aktivita. Tato aktivita vzniká za pomoci akčního potenciálu (obr. 2), jehož vznik je popsán v podkapitole 2.3. Prevodní systém srdeční je tvořen specializovanou svalovou tkání schopnou tvorby a transmise vzruchů. [1]

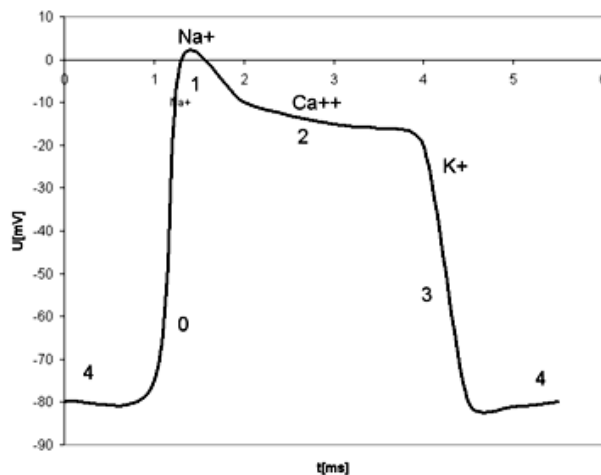
K frekvenčně nejrychlejší spontánní elektrické aktivitě dochází v sinoatriálním uzlu, umístěném v pravé síni. Zde vzniká tzv. sinusový rytmus. Po síňové svalovině jsou vzruchy dále šířeny do atrioventrikulárního uzlu, nacházejícího se vespod pravé komory blízko septa. Na komory je vzruch dále šířen za pomoci Hisova svazku, umístěného v mezikomorové přepážce. Ten je rozdělen na Tawarova raménka, která se v srdečním apexu rozvětvují na Purkyňova vlákna, která vedou vzruch k bázím komor. [1]



Obrázek 1 - Prevodní systém srdeční [13]

Je-li stimulace dostatečně silná (impulz je dostatečně silný, prahový) reaguje srdce systolickým stahem. Pokud stimulace nedosáhne prahové hodnoty, žádná akce nenastává dle zákona „všechno nebo nic“. [1]

Jsou-li ionty rovnoměrné koncentrace ve vnitrobuněčné i mimobuněčné tekutině, vzniká takzvaný klidový membránový potenciál, jehož hodnota se pohybuje přibližně od -70 až -90mV. Hlavní roli v tomto systému hraje nejvíce zastoupený intracelulární iont K^+ a extracelulární iont Na^+ . [1]

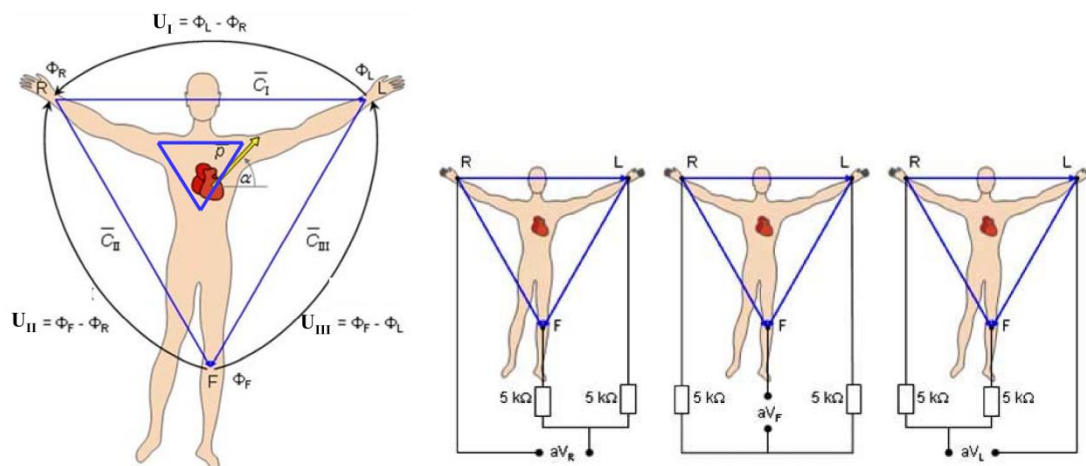


Obrázek 2 - Vznik akčního potenciálu v buňce myokardu [12]

2.2 Snímání EKG

EKG (elektrokardiogram) slouží ke snímání potenciálových změn v srdci. Elektrody EKG se lepí přímo na kůži těla. Jako EKG svod označujeme rozdíl napětíových potenciálů mezi dvojicí elektrod. Snímání EKG a hodnocení EKG je prováděno nejčastěji 12 elektrokardiografickými svody, které slouží ke snímání potenciálů mezi různými částmi těla. [6]

Princip tohoto měření vychází z tzv. Einthovenova trojúhelníku (obr. 3), který je tvořen třemi body. Dva z nich jsou umístěny přibližně na ramenou (na levém a na pravém), třetí je umístěn pod bránicí. Tento princip pak umožňuje umístění elektrod na končetiny, kdy levá a pravá horní končetina (první a druhá elektroda) má přibližně stejný potenciál, jako jsou potenciály ramen, a potenciál levé dolní končetiny (třetí elektroda) je přibližně stejný, jako potenciál bránice (přibližně nulový). [6]



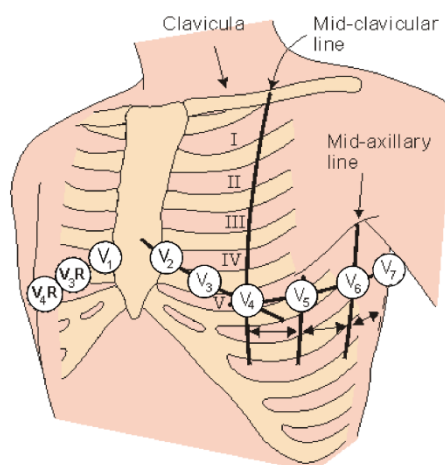
Obrázek 3 - Vlevo: Einthovenovy bipolární končetinové svody I, II a III; vpravo: unipolární Goldbergerovy svody aVR , aVF a aVL [6]

Standartní EKG o 12 svodech (obr. 4) :

Bipolární Einthovenovy svody (I,II,III) – měří napětí mezi levou rukou, pravou rukou a levou nohou [6]

Augmentované unipolární končetinové svody (aVF , aVL a aVR) – využívá se stejných elektrod jako u bipolárních svodů s tím rozdílem, že u unipolárních svodů je napětí definováno mezi jedním z vrcholů trojúhelníku a průměrem zbývajících dvou vrcholů. [7]

Unipolární hrudní svody (V1- V6) – Tyto elektrody slouží k měření potenciálů nejvíce ovlivňovaných srdcem a právě z tohoto důvodu jsou umístěny v jeho nejbližším okolí. Napětí je měřeno proti Wilsonově svorce s přibližně nulovým potenciálem. [7]



Obrázek 4 - Umístění hrudních svodů [6]

Obvykle je EKG snímáno jen osmi svody (I, II, V1 až V6), zbylá 4 napětí se dají dopočítat. [7]

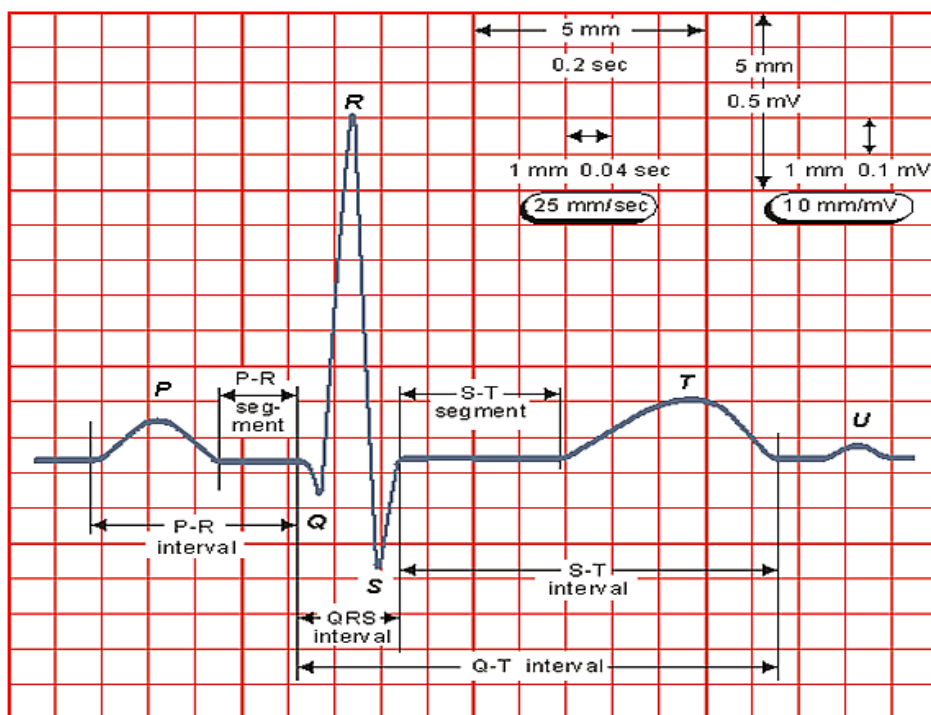
Ortogonalní svody – jedná se o tři navzájem kolmé svody x, y a z. Vektor popisující směr srdeční aktivity (jeho koncový bod) vytváří tzv. vektorkardiogram. Ten obsahuje informace o srdeční aktivitě bez redundantních informací. Užívání a hodnocení těchto svodů však není příliš časté. [7]

2.3 Vlny v EKG

Dojde-li k pohybu Na^+ iontů z extra do intracelulárního prostředí, polarizační napětí stoupá z předchozích -70 až 90mV k 0mV až +20mV (vzniká takzvaný akční potenciál) a dochází k depolarizaci. Během depolarizace se buňka 300ms nachází v takzvané refrakterní fázi, kdy je imunní proti dalšímu podráždění. Fáze depolarizace je na EKG znázorněna vlnou P (šíření depolarizace po síních) a komplexem QRS (obr. 5), který představuje postupnou depolarizaci mezikomorové přepážky, apexu a následně i srdečních bází. Po depolarizace přichází vlivem vstupování Ca^{2+} do buňky přibližně 200 – 300ms trvající fáze „plató“. Ta je na EKG znázorněna úsekem ST. Následuje repolarizace, která je následkem aktivity ATP poháněnou sodíko-draselnou pumpou, která vrací ionty do svých původních prostředí a vyrovnává koncentrace iontů. Repolarizace je na EKG znázorněna vlnou T, která odpovídá postupné repolarizaci komor. Vznik akčního napětí na buňce myokardu je i s vlivem jednotlivých iontů znázorněn na (obr. 2). V následující tabulce (tab. 1) je shrnutí trvání jednotlivých vln a komplexů EKG křivky a původ jejich vzniku. [1]

Tabulka 1 - Trvání jednotlivých vln a komplexů EKG křivky a původ jejich vzniku [1]

Vlna (komplex)	Doba trvání [ms]	Původ vzniku
P	80-100	sinoatriální uzel
PQ	80-100	zpomalení v AV uzlu
QRS	80-100	postupná depolarizace komor
ST	350-400	repolarizace komor



Obrázek 5 - Základní vlny a intervaly v EKG [6]

2.4 Poruchy srdečního rytmu

Pokud dojde ke změně automacie tzv. pacemakeru (SA uzel), dochází ke generování vzruchů s vyšší nebo nižší frekvencí. Pokud dojde k rychlejšímu generování vzruchů, dochází k tzv. tachykardii, pokud je generování zpomalené, vzniká bradykardie. [2]

Pokud je vzruch generován z místa, ve kterém ke generování vzruchu docházet nemá (nefungují za normálních okolností jako pacemakery), nazýváme tyto místa jako ektopická ložiska. Ty mohou vznikat následkem např. srdeční ischemie. [2]

Pokud je vedení v některém místě porušeno, nebo prodlouženo, mluvíme o srdeční blokádě. [2]

Pokud dojde k ischemii v ohraničené oblasti myokardu, časy převodu v tkáni se začnou různit a jako následek mohou vznikat tzv. krouživé vzruchy nebo např. reentry, na jehož podkladu vzniká fibrilace síní nebo komor. [2]

2.4.1 Některé typy arytmií

Sínový uniklý stah - Nejedná se o primární onemocnění, spíše se jedná o sekundární příznak poruchy vedení nebo poruchy depolarizace. Po sinusovém stahu selhává depolarizace v SA

uzlu. Je viditelná abnormální vlna P, z důvodu začátku excitace v jiném místě, než v SA-uzlu. Po této abnormalitě se stahy vrací k sinusovému rytmu. [3]

Nodální (junkční) uniklý stah – Při převaze aktivity AV-uzlu se rytmus nazývá nodální – junkční uniklý stah. V tahových stazích se nevyskytují vlny P z důvodu nepřítomnosti síňové kontrakce. QRS komplex je dále normální. [3]

Komorový uniklý stah – Vzniká následkem výpadku SA-uzlu a po krátké pauze se objevuje abnormální QRS a abnormální vlna T. Komorové ložisko pak po dobu jednoho stahu přebírá kontrolu nad srdeční frekvencí, následně se rytmus vrací k sinusovému rytmu. [3]

Komorová extrasystola – Vznikají při brzké depolarizaci komor. Komplex QRS se jeví jako abnormální. Při nízké četnosti nejsou komorové extrasystoly život ohrožující, avšak pokud se komorová extrasystola objeví v časně fázi vlny T, může srdce snadno přejít k fibrilaci. Na EKG je QRS komplex širší, a vlna T se také jeví jako abnormální. Může vznikat i takzvaný fenomén R na T, pokud komorová extrasystola nasedá na vrcholek vlny T předchozího stahu. Vlna P může chybět, další pak přichází v očekávaném okamžiku. [3]

Supraventrikulární extrasystola – Činí výskyt cyklu P vlny nepravidelným. Většinou se vlna P nevyskytuje v očekávané době, ale objevuje se až po kompenzační pauze. [3]

Síňová extrasystola – Vzniká při brzké depolarizaci síní. Má abnormální vlnu P, ale komplexy QRS mají stejný tvar jako při sinusovém rytmu. [3]

Junkční extrasystola – P vlna se u junkčních extrasystol vyskytovat vůbec nemusí, nebo se vlna P může objevit těsně před normálně tvarovaným komplexem QRS nebo například těsně po tomto komplexu. QRS komplex je stejně jako u síňových extrasystol tvarován jako sinusový rytmus. [3]

Síňová tachykardie – Dochází ke stahování síní ve frekvenci rychlejší než 150 stahů/min. Vlny P mohou být položeny na vlny T předešlého stahu. Tvar komplexu QRS a vlny T je většinou stejný jako u sinusového rytmu. [3]

Síňový flutter – Pokud je frekvence vyšší než 200 stahů/min. , AV-uzel není schopen takto vysokou frekvenci převádět. Dochází ke vzniku AV-blokády a pokud frekvence překročí 250 stahů/min., hovoříme o síňovém flutteru. V tomto případě není na EKG mezi vlnami izoelektrická linie. [3]

Komorová tachykardie – Komplexy QRS bývají v případě této tachykardie defektní (šíření abnormální cestou), široké, vlny P nemusí být přítomny. Tvar QRS je proměnný a nepravidelný. Jako komorovou tachykardii nazýváme moment, kdy se ložisko v komorové svalovině depolarizuje s vysokou frekvencí. Na EKG je komorová tachykardie vidět jako množství po sobě jdoucích komorových extrasystol. [3]

Fibrilace síní – Vyskytují se ve formě paroxysmální nebo chronické. Jedná se o nejčastější tachyarytmii. Výskyt výrazně stoupá s věkem [4]. Vzniká nezávislými stahy jednotlivých svalových vláken v srdci. Na EKG se důsledkem nezávislých stahů neobjevují vlny P, objevují se pouze nepravidelné křivky. Dochází k „bombardování“ AV-uzlu depolarizačními vlnami. [3]

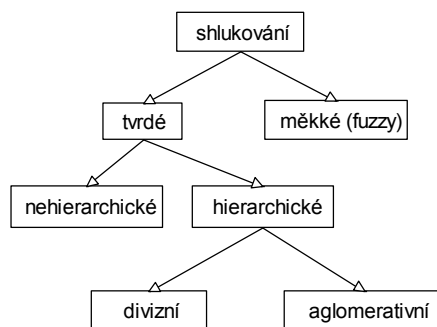
3. Shluková analýza

3.1 Úvod do shlukové analýzy

Úkolem shlukové analýzy, která bude použita v této práci, je roztřídit objekty (v tomto případě budeme chápat pojem objekty jako jednotlivé komplexy vln v EKG signálu) do shluků, ve kterých jsou si objekty navzájem podobné a bez podobností nebo s co nejmenší podobností s objekty ve shlucích ostatních. [9]

Se shlukovou analýzou se běžně setkáváme v běžném životě, pokud potřebujeme určité objekty nebo situace klasifikovat do určitých skupin. Metody shlukové analýzy jsou známy už poměrně dlouho, v poslední době ale zažívají renesanci díky metodám automatického sběru dat. Jejich využití je významné v oblasti expertních systémů a nebo právě u analýz signálu, podobných, jakými se bude zabývat tato práce. [10]

Shlukování můžeme dělit (obr. 6) na měkké (fuzzy) a tvrdé. Měkké shlukování je pro tuto práci poměrně zbytečné, protože je třeba klasifikovat shluky přesně, ne přibližně, jak vybízí použití fuzzy shlukování. Tvrdé shlukování dělíme dále na hierarchické, jež má více možných řešení a k jehož interpretaci využíváme dendrogramů a na nehierarchické, které nabízí rozklad na n shluků. Hierarchické shlukování se v poslední řadě dělí na aglomerativní, přičemž využíváme jednotlivých objektů a ty spojujeme do shluků, a na divizní, kdy vycházíme z jednoho shluku jako celku a ten dále dělíme do shluků menších. [6]



Obrázek 6 - Rozdělení shlukové analýzy [14]

V této práci se uplatní metody, které jsou založené na tvrdém, hierarchickém, aglomerativním shlukování, kdy se předpokládá, že signál obsahuje objekty (komplexy) jako samostatné, a shlukovou analýzou budou objekty děleny do jednotlivých shluků.

Mimořádnou pozornost je nutno v tomto případě věnovat výběru příznaků jednotlivých objektů. U jednotlivých příznaků je vyžadováno dbát na jednotky, ve kterých jsou dané příznaky uváděny. Pokud bude použito více než jednoho příznaku, je nutné příznaky pomocí vhodné metody standardizovat. Standardizace je nutná hlavně v případě, kdy jsou jednotky, ve kterých jsou příznaky uváděny, zaměnitelné za jiné. To by se v případě změny jednotek mohlo projevit již v případě, že by byl proveden „scatter“ plot těchto hodnot. Shluky hodnot by v tomto případě mohly vypadat úplně jinak, než v případě z předchozími jednotkami. Standardizace je tedy nutná z důvodu zamezení nepříjemných zkreslení, která by ovlivnila celkový výsledek analýzy. [8]

3.2 Postup u aglomerativních metod shlukové analýzy

3.2.1 Získání matice dat

Tato matice bude obsahovat v řádcích jednotlivé objekty, kdežto ve sloupcích budou jednotlivé příznaky pro vybrané objekty. V případě objektů můžeme hovořit například o subjektech léčby, a v případě příznaků (atributů) budeme mluvit například o reakci subjektu na určitý typ léčby, vyjádřenou číselně. Čím více příznaků bude vybráno pro metodu shlukové analýzy, tím by měla analýza obsahovat buď více skupin, a nebo by se měla zvýšit její celková přesnost za předpokladu, že jeden z příznaků přímo ovlivňuje příznaky další. [9]

3.2.2 Standardizace dat

Data v získané matici dat mohou být v různých jednotkách. To může způsobovat zkreslení, jak je zmíněno výše, už při „scatter“ plotu dat. Představme si například, že jsme změřili výšku několika objektů v centimetrech, a několika objektů v palcích. V tomto případě potřebujeme dané jednotky převést na nové, univerzální, bezrozměrné veličiny. K tomu nám slouží metody standardizace dat.[9]

Pro výpočet standardizace dat potřebujeme nejdříve vypočítat průměr m_f pro dané proměnné $f(1)$.

$$m_f = \frac{1}{n} (x_{1f} + x_{2f} + \dots + x_{nf}). \quad (1)$$

Dále potřebujeme vypočítat střední odchylku. Pro tyto účely můžeme buď použít klasický výpočet pro střední odchylku (2), a nebo výpočet pro střední absolutní odchylku (3), která nám vyloučí případ, že by špatně naměřená (příliš malé nebo příliš vysoké hodnoty) data zkreslila, díky použití umocnění, celý výpočet.

$$std_f = \frac{1}{n-1} \{(x_{1f} - m_f)^2 + (x_{2f} - m_f)^2 + \dots + (x_{nf} - m_f)^2\}. \quad (2)$$

$$s_f = \frac{1}{n} \{|x_{1f} - m_f| + |x_{2f} - m_f| + \dots + |x_{nf} - m_f|\}. \quad (3)$$

Získané hodnoty pak dále použijeme pro samotný výpočet pro konkrétní standardizovaná data z_{if} (4).

$$z_{if} = \frac{x_{if} - m_f}{s_f}. \quad (4)$$

Tímto výpočtem získáme novou, standardizovanou matici dat, kterou můžeme použít pro další postup ve výpočtu. [8]

3.2.3 Výpočet matice podobností

Tento krok slouží k výpočtu jednotlivých vzdáleností mezi jednotlivými páry ať už standardizovaných nebo přirozených objektů. Nejčastěji se k tomuto kroku využívá výpočet euklidovských distancí. Pokud máme například matici, která obsahuje pět objektů, pomocí výpočtu podobností získáme deset hodnot podobností mezi objekty, které ukládáme to tzv. podobností matice.

Vždy máme dva typy podobnostních koeficientů :

Koeficient podobnosti – čím větší je jeho hodnota, tím jsou si objekty bližší.

Koeficient nepodobnosti – čím menší je jeho hodnota, tím jsou si objekty bližší.

Nejčastěji tyto koeficienty získáváme výpočtem z dvojdimenzionálního prostoru pomocí tzv. Euklidovských distancí (5). [9]

$$e_{jk} = \sum_{i=1}^n (x_{ij} - x_{ik})^2. \quad (5)$$

3.3 Provedení vlastní metody shlukové analýzy

Jakmile jsme získali matici podobností, můžeme začít provádět samotnou shlukovou analýzu. K tomuto kroku můžeme využít hned několik metod analýzy, kdy každá z nich využívá jiný způsob vyhledávání podobností mezi objekty. Představíme si tedy tyto čtyři metody:

1. UPGMA

Z angličtiny *unweighted pair-group method using arithmetic averages*, přeloženo jako neváhovaná metoda párů používající aritmetické průměry, je jedna z nejpoužívanějších metod shlukové analýzy. Pro použití této metody používáme matici podobností, spočítanou v předchozích krocích.

Pro výpočet metody je vždy základní najít nejmenší hodnotu euklidovské vzdálenosti mezi objekty, a následně sloučit shluky, které tuto hodnotu protínají. V prezentovaném příkladu (tab. 2) se jedná o shluky 3 a 4.

Tabulka 2 - Výběr nejmenší hodnoty podobnosti v matici podobností [9]

-	1	2	3	4	5
1	-	-	-	-	-
2	18,0	-	-	-	-
3	20,6	14,1	-	-	-
4	22,4	11,2	5,00	-	-
5	7,07	18,0	25,0	25,5	-

Další hodnoty matice jsou pak vypočítány pomocí aritmetických průměrů jednotlivých objektů s dalšími objekty v ostatním slucích (6).

$$\begin{aligned}
 m_{1(3,4)} &= \frac{1}{2}(e_{13} + e_{14}). \\
 m_{2(3,4)} &= \frac{1}{2}(e_{23} + e_{24}). \\
 m_{5(3,4)} &= \frac{1}{2}(e_{53} + e_{54}).
 \end{aligned}
 \tag{6}$$

Následně provedeme ještě sloučení shluků 1 a 5, čímž dostaneme následující matici podobností (tab. 3) :

Tabulka 3 - Nově získaná matice podobností [9]

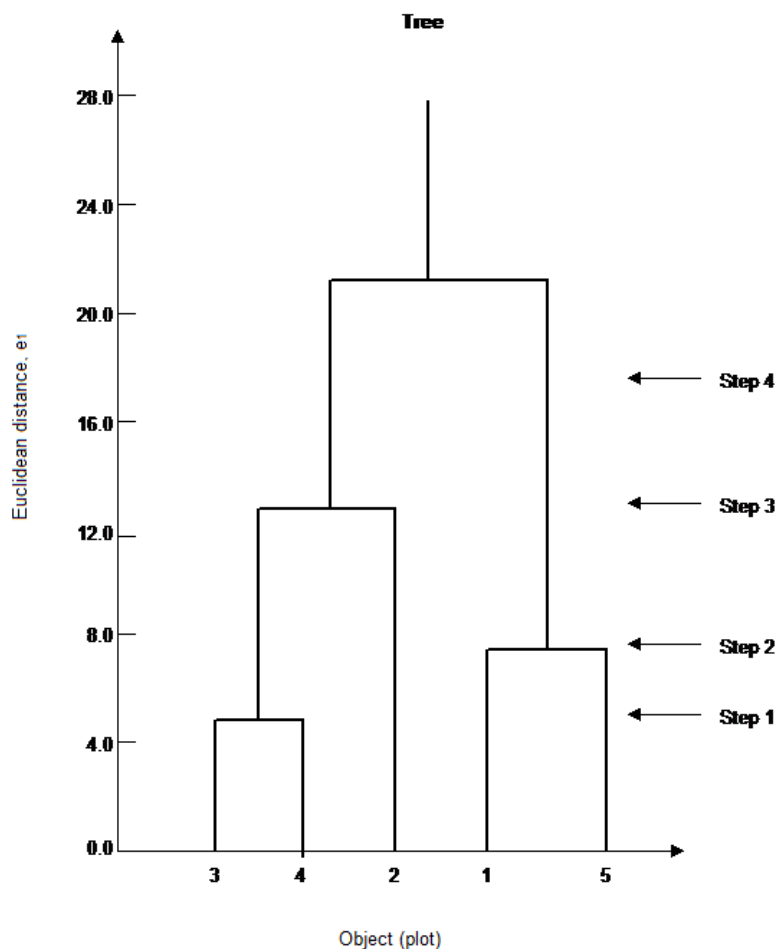
-	2	(34)	(15)
2	-	-	-
(34)	12,7	-	-
(15)	18,0	23,4	-

A nakonec vybereme nejmenší hodnotu i z této matice, provedeme sloučení opět dle principů (6), a získáme výslednou matici euklidovských distancí (tab. 4). V této matici už nemáme shluky, které bychom mohli sloučit.

Tabulka 4 - Výsledná matice podobností [9]

-	(15)	(234)
(15)	-	-
(234)	21,6	-

Díky získaným novým shlukům, a nejmenším hodnotám euklidovských distancí můžeme sestavit výsledný dendrogram (obr. 7). Ten na ose x nese popisky jednotlivých objektů, a na ose y jsou euklidovské distance.



Obrázek 7 - Výsledný dendrogram [9]

2. SLINK

Metoda SLINK (*single linkage clustering method*) pracuje na podobném principu jako výše zmíněná metoda UPGMA s tím rozdílem, že pokud v matici euklidovských distancí nalezneme nejmenší hodnotu, další shluky a další hodnoty distancí již nepočítáme pomocí aritmetického průměru, ale vezmeme obě hodnoty, které chceme sloučit a vybereme z nich tu nejmenší.[9]

$$e_{n(ij)} = \min(e_{ni}; e_{nj}). \quad (7)$$

3. CLINK

CLINK (*complete linkage clustering method*) funguje obdobně jako metoda SLINK, jen s tím rozdílem, že oproti minimální hodnotě euklidovské distance vybíráme naopak hodnotu maximální.[9]

$$e_{n(ij)} = \max(e_{ni}; e_{nj}). \quad (8)$$

4. Wardova shlukovací metoda minimálního rozptylu

Hned po metodě UPGMA je Wardova shlukovací metoda nejpoužívanější metodou pro shlukovou analýzu. Tato metoda oproti třem předchozím nepoužívá pro výpočet shluků matici podobností, ale je možno použít hned počáteční matici objektů a jejich atributů.

Jako první musíme vypočítat hodnotu E (9), což je v tomto případě suma všech shluků, pro kombinaci všech atributů všech možných objektů.[9]

Tabulka 5 - Vstupní data [9]

Objekty	1	2	3	4	5
Atribut 1	10	20	30	30	5
Atribut 2	5	20	10	15	10

Budeme li tedy vycházet ze vstupních dat (tab.5), bude výpočet vypadat takto(9) :

$$E = \underbrace{(10 - 10)^2 + (5 - 5)^2}_{\text{Shluk 1}} + \underbrace{(20 - 20)^2 + (20 - 20)^2}_{\text{Shluk 2}} + \underbrace{(30 - 30)^2 + (10 - 10)^2}_{\text{Shluk 3}} + \underbrace{(30 - 30)^2 + (15 - 15)^2}_{\text{Shluk 4}} + \underbrace{(5 - 5)^2 + (10 - 10)^2}_{\text{Shluk 5}} = 0.0. \quad (9)$$

Dále provedeme všechna možná sloučení objektů a jejich atributů do shluků po maximálně dvou objektech. Říkejme tomuto slučování krok 1. Toto sloučení počítáme vypočtením průměru z hodnot atributů. Pro sloučení hodnot pro objekty 1 a 2 ze vstupních dat (tab.5), bude výpočet vypadat takto (10) : [9]

$$\underbrace{(12) = \frac{10 + 20}{2} = 15,}_{\text{Pro atribut 1}} \quad \underbrace{(12) = \frac{5 + 20}{2} = 12.5.}_{\text{Pro atribut 2}} \quad (10)$$

Po-té, co provedeme výpočet všech E pro všechna možná sloučení hodnot do shluků o dvou objektech, začneme slučovat i další objekty do dalších, menších shluků, např. (134) (2) a (5), nebo (12), (34) a (5) a opět provedeme výpočet všech E. Tomuto kroku říkáme krok 2.

V kroku 3 opět sloučíme již vzniklé shluky do shluků ještě obsáhlejších a vypočteme E a nakonec provedeme sloučení v tomto případě vzniklých shluků (125) a (34) a provedeme výpočet E, čímž dostaneme nejvyšší hodnotu větve v dendrogramu. Tento krok nazvěme tedy krokem 4. Dostaneme tedy hodnoty znázorněné v obrázku (obr. 8). [9]

Step	Possible partitions				E
1	(12)	3	4	5	162.5
	(13)	2	4	5	212.5
	(14)	2	3	5	250.
	(15)	2	3	4	25.
	(23)	1	4	5	100.
	(24)	1	3	5	62.5
	(25)	1	3	4	162.5
	(34)	1	2	5	12.5*
	(35)	1	2	4	312.5
	(45)	1	2	3	325.0
2	(34)	(12)	5		175.0
	(34)	(15)	2		37.5*
	(34)	(25)	1		175.
	(134)	2	5		316.7
	(234)	1	5		116.7
	(345)	1	2		433.3
3	(234)	(15)			141.7*
	(125)	(34)			245.9
	(1345)	2			568.8
4	(12345)				650.

Obrázek 8 - Vytvoření všech možných shluků a výpočet jejich hodnot E [9]

Dendrogram opět sestavujeme pouhým vybráním vždy nejmenší hodnoty daného shlukovacího kroku. Hodnoty E se nám během shlukování nelineárně zvyšují. Nakonec provedeme sestavení samotného dendrogramu. [9]

3.4 Výpočet Pearsonova korelačního koeficientu

Pearsonův korelační koeficient počítáme proto, abychom zjistili, jakým způsobem nám koreluje hodnoty atributů mezi jednotlivými objekty. Pokud si jednotlivé atributy označíme jako x a y , bude výpočet vypadat takto (11) :

$$r_{(x,y)} = \frac{\sum xy - \frac{\sum x \sum y}{N}}{\sqrt{\left(\sum x^2 - \frac{(\sum x)^2}{N}\right) \left(\sum y^2 - \frac{(\sum y)^2}{N}\right)}}. \quad (11)$$

Dokonalá shoda znamená hodnotu $r_{(xy)}=1$, neshodě odpovídá $r_{(xy)}=0$ a negativní korelaci odpovídá hodnota $r_{(xy)}=-1$. Pokud $r_{(xy)}\geq 0,8$, je zkrácení dendrogramem považováno za přijatelné. [7]

4. Vybrané příznaky a jejich zdůvodnění

Úkolem práce bylo, vybrat nejvhodnější příznaky pro automatickou shlukovou analýzu normálních a abnormálních QRS komplexů v EKG signálu. K získání základních parametrů z EKG signálu (detekce a jeho rozměření) bylo použito programu z disertační práce Ing. Martina Vítka, Ph.D [15]. Jednotlivé vlny jsou detekovány pomocí vlnkové transformace. Například na obrázku (obr. 10) je vidět ono rozměření, kde vlna Q je zaznačena zeleně, vlna R červeně a vlna S fialově.

4.1 Tvarový faktor

Jako první atribut, který je použitelný k popisu QRS komplexu, je možno použít tzv. **tvarový faktor** (*angl. Form factor*). Pro získání tvarového faktoru potřebujeme několik parametrů úseku signálu. První z nich je aktivita, což je v podstatě hodnota rozptylu σ_x^2 pro daný úsek signálu $x(n)$ a druhým parametrem je mobilita $M(x)$ (12), kterou počítáme jako odmocninu podílu aktivity první derivace signálu $\sigma_{x'}^2$ aktivitou originálního signálu σ_x^2 : [11]

$$M(x) = \left[\frac{\sigma_{x'}^2}{\sigma_x^2} \right]^{\frac{1}{2}} = \frac{\sigma_{x'}}{\sigma_x}. \quad (12)$$

Třetím parametrem je komplexita (13), nebo už tzv. samotný tvarový faktor. Ten je vypočítán jako podíl mobility první derivace signálu k mobilitě originálního signálu a kde $\sigma_{x''}$ je aktivita druhé derivace úseku signálu $x(n)$. [11]

$$FF = \frac{M_{x'}}{M_x} = \frac{\sigma_{x''}/\sigma_{x'}}{\sigma_{x'}/\sigma_x}. \quad (13)$$

V případě této práce je úsek $x(n)$ definován jako délka trvání cyklu od začátku QRS komplexu po konec T vlny.

4.2 Vzdálenostní faktor

Po rozměření signálu a výpočtu jednotlivých vzdáleností mezi vlnami vzniklo mnoho hodnot vzdáleností. Každou z nich bylo možno zvolit jakožto příznak, protože většina z nich se lišila právě v místech výskytu abnormálního komplexu. Proto bylo použito abstraktního **vzdálenostního faktoru** (14), obdoby již zmíněného tvarového faktoru, jehož výpočet vytvořil sám autor práce. Hodnoty jsou počítány pomocí vzdáleností mezi vlnami QR, RS, QT a vzdáleností RR. Hodnota vzdálenostního faktoru bývá v abnormálním QRS komplexu v důsledku abnormalit v trvání celého komplexu, abnormality vlny T nebo abnormality QRS úseku, až dvakrát vyšší, než u normálního QRS komplexu.

$$VF = \frac{(QR * RS) * QT}{RR}. \quad (14)$$

4.3 R-R interval

Dalším základním příznakem při detekci arytmií je **RR vzdálenost**, tj. vzdálenost mezi dvěma vlnami R v signálu. Tato vzdálenost dobře popisuje valnou většinu typů arytmií hlavně proto, že valná většina patologických cyklů má rychlejší nástup, což vede ke zkracování R-R intervalu. První QRS komplex je ve všech analýzách v důsledku sestavení stejně dlouhých vektorů atributů vynechán, ačkoliv R vlna z prvního QRS komplexu, k jejíž detekci dojde vždy, je využita pro výpočet R-R intervalu jakožto atributu, i když není v obrázcích s detekcemi vyznačena.

4.4 Plocha pod křivkou QRS komplexu

Posledním základním příznakem, na němž je založena celá pozdější shluková analýza, jsou **plochy pod křivkou QRS komplexů**. Každý QRS komplex získává díky svému rozkmitu určitou velikost plochy pod křivkou, získanou integrací. Atribut ploch ze signálu je tedy získán rozměření signálu, vyjmutím úseku QRS komplexu, přepočtem tohoto úseku signálu do absolutních hodnot, a následnou integrací tohoto úseku lichoběžníkovou metodou. Pro co nejmenší ztrátu informace o ploše komplexů v signálu nebyl signál podroben filtraci.

5. Obsah a obsluha programového řešení

Programové řešení obsahuje tyto funkce :

- | | |
|-------------------------|--|
| - shlukovani.m | – jedná se o hlavní funkci pro celou analýzu |
| - atributy.m | – podfunkce pro srovnání atributů do matice |
| - EKGrozmereni_global.m | – funkce pro rozměření signálu [15] |
| - FF.m | – funkce pro výpočet tvarového faktoru |
| - automat.m | – funkce pro srovnání vektorů detekcí |
| - shlukovani_w117.m | – hlavní funkce v úpravě pro signál w117.mat |

Obsluha programu shlukovani.m

Pro spuštění funkce je třeba zadat například příkaz:

```
shlukovani('w075.mat',1,2)
```

Tento příkaz provede analýzu signálu w075.mat z pseudoortogonálních svodů II, V2 a V6 (jsou pevně nastaveny) za pomoci atributů R-R intervalu (1) a plochy QRS komplexů (2).

Jsou-li vybrány duplicitní atributy, program se nespustí. Spektrum signálu je upraveno vynulováním prvních šesti a posledních pěti spektrálních čar, čímž je signál vyrovnán. Pro signál w034.mat je nutné, jak je popsáno v kapitole 6.3, změnit svod, ze kterého je prováděno rozměření, na svod č. 4. To je možné na řádce 43, kdy je nutné místo proměnné x nastavit proměnnou x(4,:).

Funkce shlukovani_w117.m má strukturu téměř totožnou s funkcí shlukovani.m, jen je upravena pouze pro analýzu signálu w117.mat. Není tedy třeba volat název signálu.

Je-li zapotřebí změnit metodu u některého z testovaných signálů, je tato možnost přítomna na řádce č. 156. Příkazy pro jednotlivé metody shlukování jsou následující:

- | | |
|---------------------------|--------------|
| UPGMA | - 'average' |
| Wardova shlukovací metoda | - 'ward' |
| CLINK | - 'complete' |
| SLINK | - 'single' |

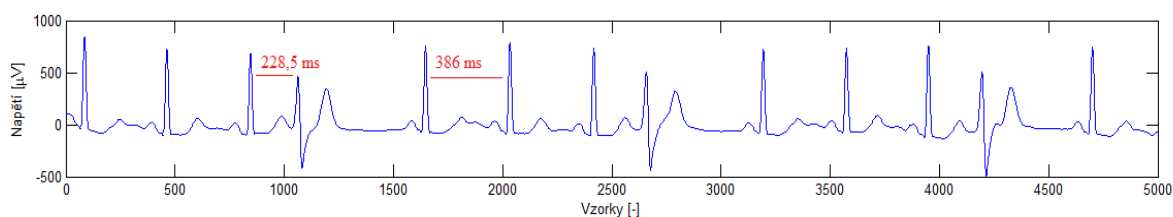
Funkce automat.m seřadí všechna rozměření, provedená pomocí funkce EKGrozmereni_global.m. Z důvodu častého vynechání některé z vln v prvním a posledním komplexu signálu jsou pomocí funkce automat.m tyto komplexy automaticky vyřazeny.

Všechny další možnosti voleb signálů či atributů jsou přítomny v nápovědách na začátku funkcí shlukovani.m a shlukovani_w117.m, včetně nápověd, jak funkce řádně spustit. Svody pro analýzu je možno v případě potřeby změnit na řádce č. 35 zapsáním do vektoru.

6. Zhodnocení výsledků práce

Ke shlukové analýze bylo v práci použito R-R intervalů vln a primárně plochy QRS křivek ze svodů II, V2 a V6. Výsledky podané analýzou stávající z R-R intervalů a ploch jsou srovnány s výsledky analýzy sestavené z atributů R-R intervalů a tvarového faktoru ze stejných svodů, jaké byly vybrány pro výpočet ploch. Svody pro plochy a tvarové faktory byly vybrány z důvodu vysokého napětového rozkmitu, který zajišťuje případnou vysokou odlišnost patologických QRS komplexů od fyziologických. Svody byly vybrány pouze tři, protože vyšší počet svodů by byl redundantní. Zároveň jsou vybrané svody pseudoortogonální [16] a je u nich tedy přítomna jistá nezávislost, která je od atributů vyžadována.

Atribut intervalu R-R vln byl vybrán z důvodu zkrácení intervalu v důsledku rychlejšího nástupu patologických komplexů, předcházející většině patologií, převážně extrasystolických. Toto časové zkrácení je jedním z nejdůležitějších dějů, které pomáhají odlišovat abnormality v EKG signálu.



Obrázek 9 - Ukázka rozdílnosti délek R-R intervalu u patologické a fyziologické vlny

Atribut plochy byl vybrán z důvodu zvýšeného časového rozšíření a většího napětového rozkmitu, případně naopak, u vln vyskytujících se v patologických QRS komplexech. Díky výše zmíněnému se při abnormalitě zvětšuje či zmenšuje plocha celého QRS komplexu oproti normálnímu sinusovému rytmu, což kladně přispívá k detekci abnormálních komplexů. Problémem atributu ploch je fakt, že plochy jednotlivých QRS komplexů se od sebe mohou výrazně lišit, proto i když jsou některé přiřazeny do samostatné skupiny, nastávají situace, kdy má jeden komplex díky výraznějšímu napětovému rozlišení plochu mnohem větší, a shluková analýza ho na základě podobnosti není schopná přiřadit k ostatním patologickým QRS komplexům v samostatném shluku.

Tvarový faktor byl vybrán jako doplňkový atribut pro třísvodovou analýzu, protože podal, stejně jako plocha QRS komplexů, poměrně dobré výsledky u jednosvodových analýz.

Program vytvořený v prostředí MATLAB 2012b poskytuje možnost výběru atributů tvarového faktoru, ploch QRS komplexu a R-R intervalů.

Téměř všechny analýzy v práci jsou vytvořeny za pomoci metody UPGMA, která poskytuje (dle Pearsonova korelačního koeficientu) nejméně zkreslené výsledky. Pro správnost některých analýz je však nutno použít i jiné metody shlukování.

Volba atributu R-R intervalu a plochy podává, co se správných analýz týče, početně podobné výsledky, jakých je dosaženo při použití atributů R-R intervalů a tvarových faktorů.

Tabulka 6 - Výsledky jednosvodové analýzy z prvního svodu signálu pro různé kombinace atributů

Signál	W006	W020	W023	W026	W028	W034	W040	W047	W054	W061	W065
RR + VF			WARD								
RR+FF											
RR+plochy			WARD								
Signál	W074	W075	W103	W105	W109	W112	W115	W117	W120	W 122	x
RR + VF				WARD							x
RR+FF								(!)			x
RR+plochy								(!)			x

Tabulka 7 - Výsledky třísvodové analýzy z pseudoortogonálních svodů pro různé kombinace atributů

Signál	W006	W020	W023	W026	W028	W034	W040	W047	W054	W061	W065
RR+FF						(!)					
RR+plochy						(!)					
Signál	W074	W075	W103	W105	W109	W112	W115	W117	W120	W122	x
RR+FF								(!)			x
RR+plochy					WARD			(!)			x

V tabulkách (tab. 6 a 7) jsou analýzy, které selhaly, vyznačeny červenou barvou. Úspěšné analýzy jsou vyznačeny zeleně, fialově jsou vyznačeny signály, u nichž byla ke správnému shlukování použita jiná metoda, než metoda UPGMA. Pokud pole obsahuje vykřičník, je signál popsán dále kvůli problematice vyhodnocení daného shlukování. Ve většině případů se jedná o manuální nastavení řezu výsledným dendrogramem. Tyto pole s vykřičníkem a s nimi spojené signály jsou popsány v kapitole 6.3.

6.1 Použití atributu R-R intervalů a ploch QRS komplexů

V této podkapitole je popsáno několik případů úspěšné a neúspěšné shlukové analýzy za pomoci atributů ploch QRS komplexů a R-R intervalů a za použití analýzy ze tří svodů signálu. Tato kombinace atributů je brána v této práci jako hlavní a směrodatná kombinace a je srovnávána s výsledky kombinace R-R intervalů s tvarovými faktory, opět ze třech pseudoortogonálních svodů signálu.

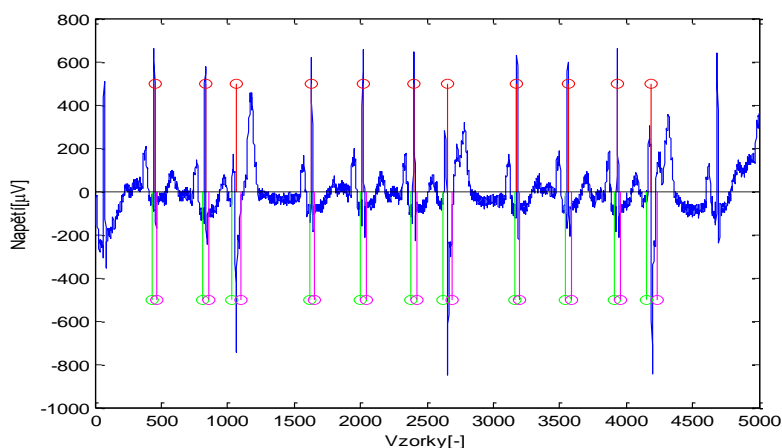
Problematika shlukové analýzy je dále popsána v závěru práce a další ukázky špatné a správné funkce tohoto typu analýzy jsou zmíněny v příloze této práce.

U všech obrázků s detekcemi jsou vždy zobrazeny jen ty komplexy, jež obsahují všechny tři požadované vlny, tedy jejich detekce je kompletní a jsou vykresleny do 2. svodu signálu.

6.1.1 Ukázky úspěšného shlukování

Signál: W075.mat

Prvním signálem, jenž demonstruje správnou funkci algoritmu shlukování za použití daných atributů, je signál w075.mat, který obsahuje 3 patologické QRS komplexy na pozicích 3,7 a 11.

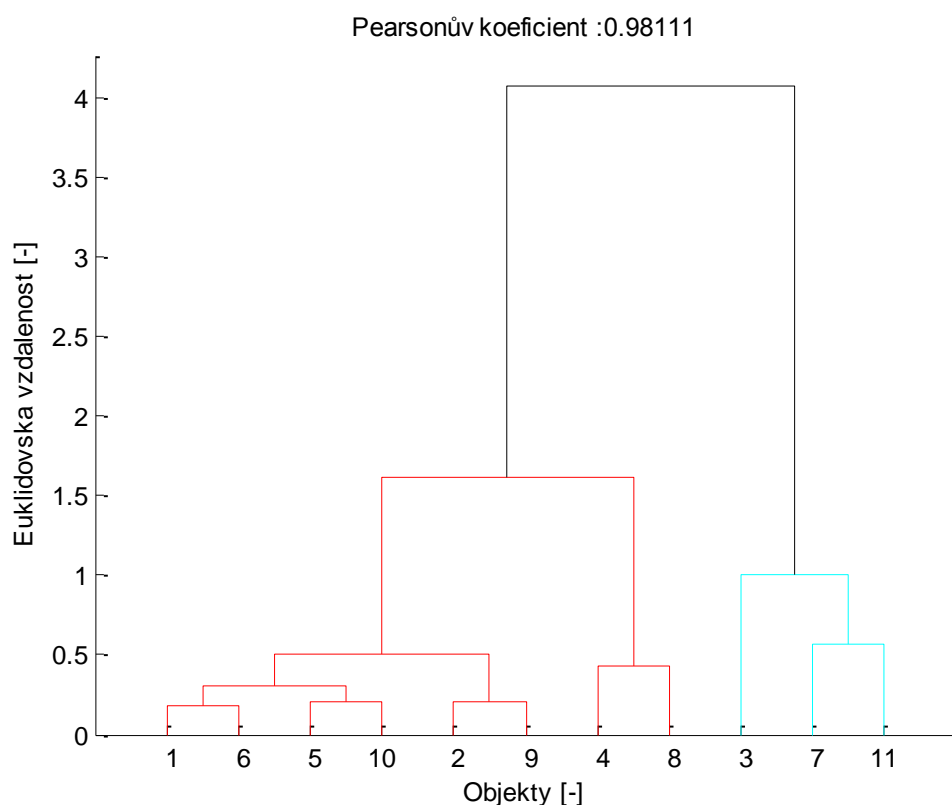


Obrázek 10 - Detekce QRS vln v signálu w075.mat

Na obrázku (obr. 10) jsou jasně vidět patologie na daných pozicích. Velká míra rozkmitu vln S a T (v případě ploch hlavně vlny S) poskytuje větší hodnotu plochy QRS komplexu, proto je plocha QRS komplexu ideálním atributem. Rozdíl ve velikosti plochy abnormálních komplexů je patrný z tabulky (tab. 8).

Tabulka 8 - Hodnoty ploch a RR intervalů pro signál w075.mat

QRS komplex	1	2	3	4	5	6	7	8	9	10	11
Plochy II	6787,5	7430,3	15329	7658,5	7608,8	7290,6	18207	7838,9	7006,8	7533,2	18888
Plochy V2	14532	13880	20279	14596	15295	14916	21750	14439	14149	15578	22901
Plochy V6	4918,3	5879,5	10289	4862	4504,6	5082	9242,2	4902,7	5486	4847,1	10227
RR interval	378	382,5	228,5	573	386	383	249,5	527,5	381,5	374	254,5

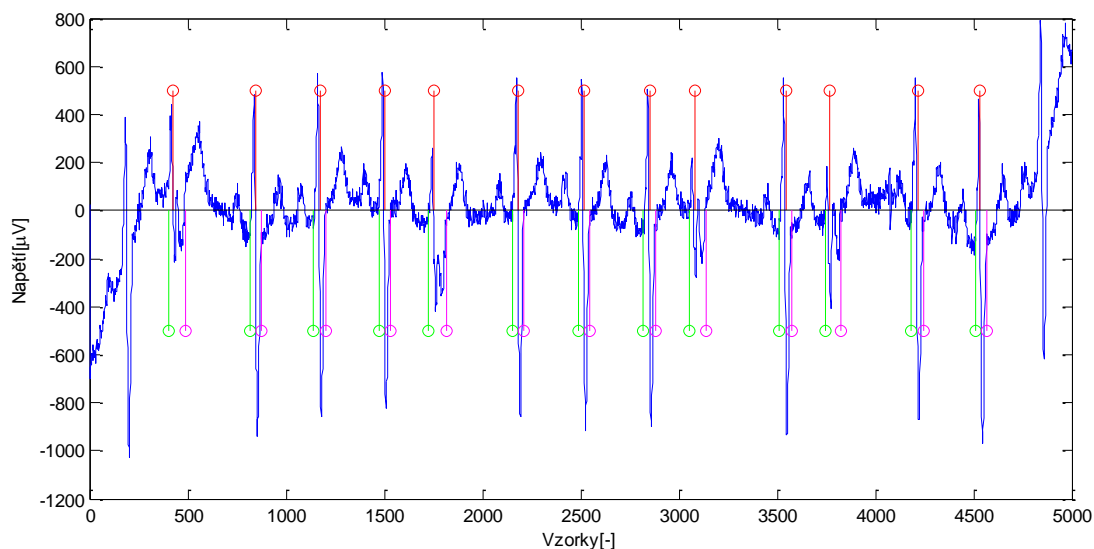


Obrázek 11 - Provedení shlukové analýzy pro signál w075.mat metodou UPGMA

V tomto případě byla shluková analýza provedena pomocí metody UPGMA. Komplexy 3, 7 a 11 jsou přiřazeny do samostatného shluku (obr. 11).

Signál: W122.mat

Jako druhý signál pro ukázkou správné funkce algoritmu byl použit signál W122.mat z knihovny signálů. Obsahuje 4 patologické QRS komplexy.



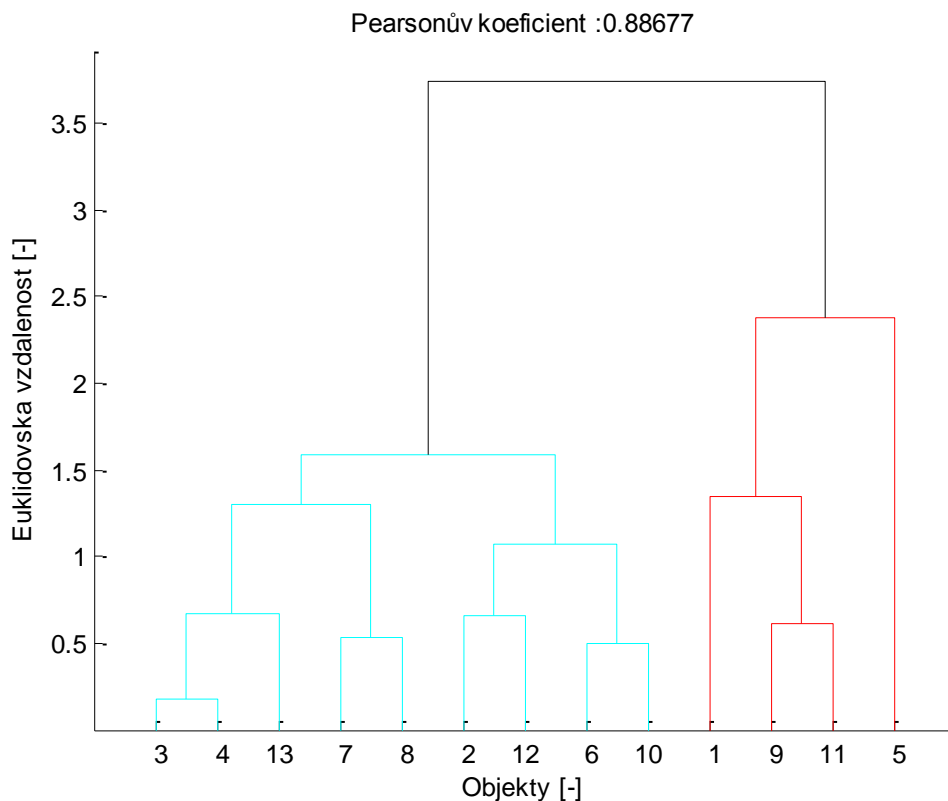
Obrázek 12 - Ukázka detekce QRS vln

Ze signálu (obr. 12) je patrné, že patologické vlny se vyskytují na pozicích 1,5,9 a 11. Jedná se s největší pravděpodobností o extrasystoly. Patrný je také menší rozkmit mezi R a S vlnou. Je také snížena amplituda R vlny. Díky tomuto faktu budou plochy těchto komplexů obecně menší, než plochy QRS fyziologických komplexů s mnohem větším rozkmitem (tab. 9).

V následujícím příkladu je použito atributu ploch QRS komplexů ze všech tří svodů a jako druhý atribut byly zvoleny R-R intervaly z prvního svodu signálu. Výsledný dendrogram byl vytvořen za pomoci shlukovací metody UPGMA.

Tabulka 9- Hodnoty ploch a RR intervalů pro signál w122.mat

QRS komplex	1	2	3	4	5	6	7	8	9	10	11	12	13
Plochy II	10973	22485	20931	21340	19433	20588	21722	21156	9694,2	21814	10090	21778	23574
Plochy V2	23954	17806	18604	18719	21934	20051	21285	20361	21197	19716	22092	18583	18402
Plochy V6	11625	19291	19374	19900	15557	19737	18266	18747	9582	19815	10907	17943	18038
RR interval	232,5	420,5	326	329	253	429	334	334	229	462	223,5	448	320



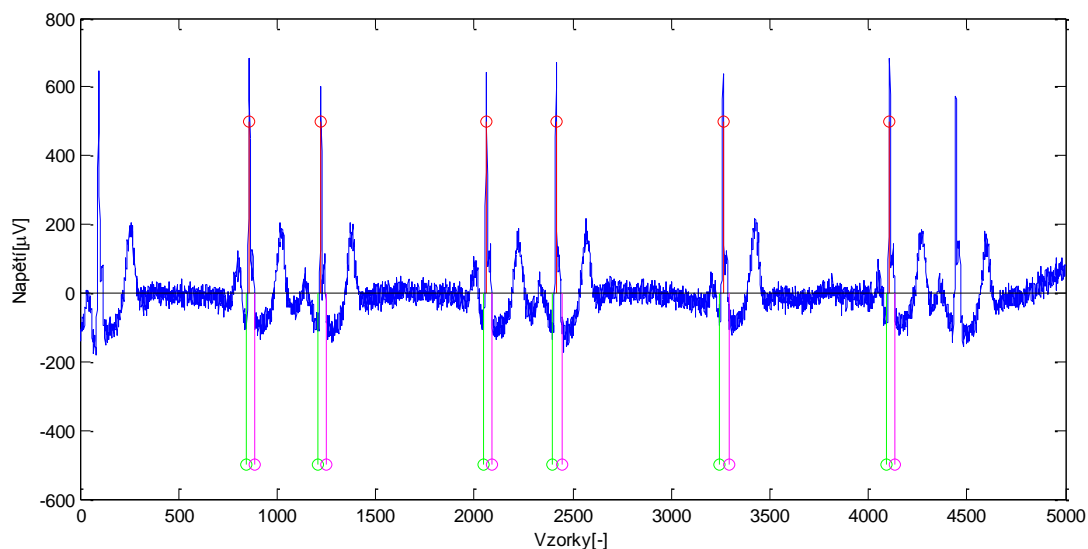
Obrázek 13 - Samotné shlukování pro signál W122.mat pomocí atributů R-R intervalu a ploch QRS komplexů

Na obrázku (obr. 13) je zřetelné, že za použití metody využívající kombinace atributů plochy a R-R intervalu lze dosáhnout správných výsledků shlukování. Metoda UPGMA v tomto případě vyhodnotila a přiřadila do správného shluku patologické vlny 1,9,5 a 11.

6.1.2 Ukázky neúspěšného shlukování

Signál: W023.mat

Tento signál obsahuje patologické vlny na pozicích 2, 4 a 7. Díky odstranění poslední vlny v důsledku úpravy vektorů detekce pro analýzu je sedmý komplex vynechán. Do samostatného shluku by se tedy měly zařadit pouze vlny 2 a 4. Vlny v signálu se nejeví patologické rozdílností v rozkmitu, pouze se zdají být extrasystolické, což vede ke zkrácení R-R intervalů (obr. 14).



Obrázek 14 - Detekce QRS komplexů v signálu w023.mat

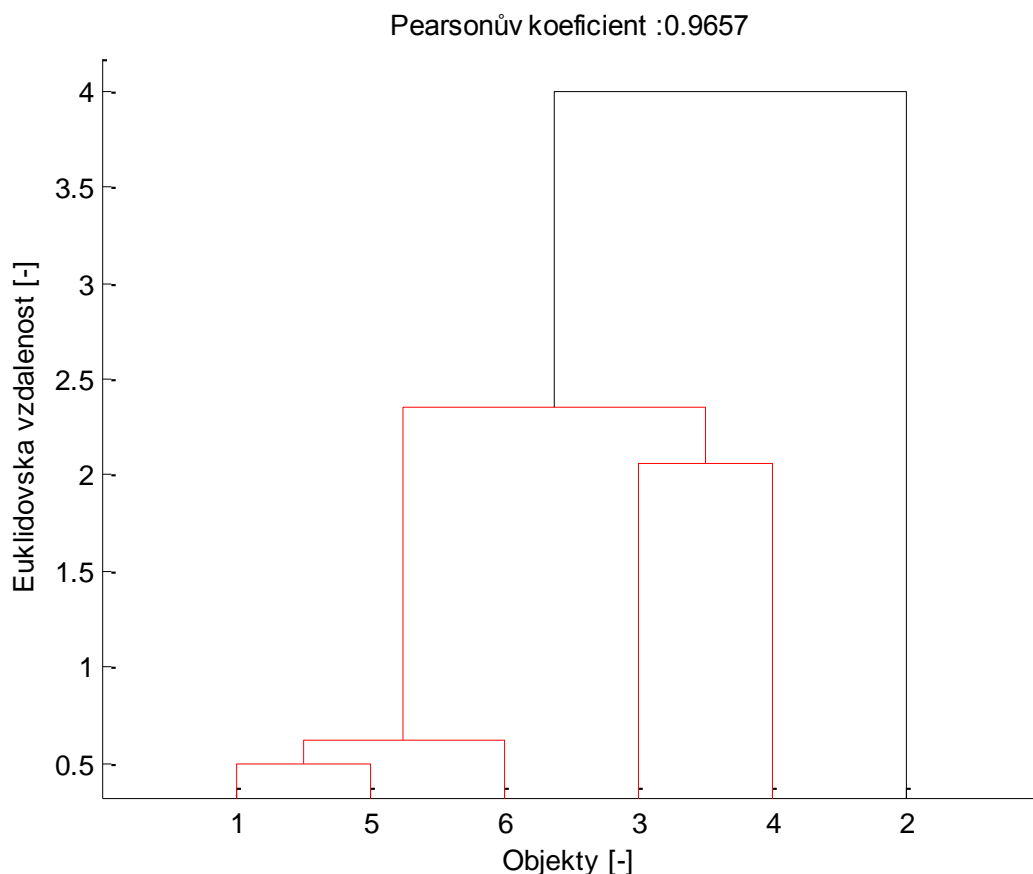
Jak je patrné z tabulky atributů (tab. 10), jsou plochy v tomto případě atributem nedostačujícím k zařazení do samostatného patologického shluku.

Tabulka 10 - Atributy signálu w023.mat

QRS komplex	1	2	3	4	5	6
Plochy II	7348,9	6197,3	7305,9	7311,9	7303	7225,4
Plochy V2	18132	17746	20551	20172	18484	17739
Plochy V6	9649,1	11386	8961,4	9269,5	9479,7	9746,5
RR interval	764,5	364	839,5	353,5	847	845

Vhodným atributem pro tuto analýzu jsou v tomto případě tedy pouze intervaly mezi R vlnami, které jsou na pozicích 2 a 4 kratší, než u ostatních QRS komplexů, jejichž intervaly jsou naopak prodlouženy v důsledku kompenzačních pauz, což může způsobit přiřazení fyziologických komplexů mezi komplexy patologické.

Shluková analýza následně vyhodnotí jako patologickou pouze vlnu č. 2 (obr. 15), jež má oproti ostatním vlnám mírně menší velikost plochy pod křivkou.

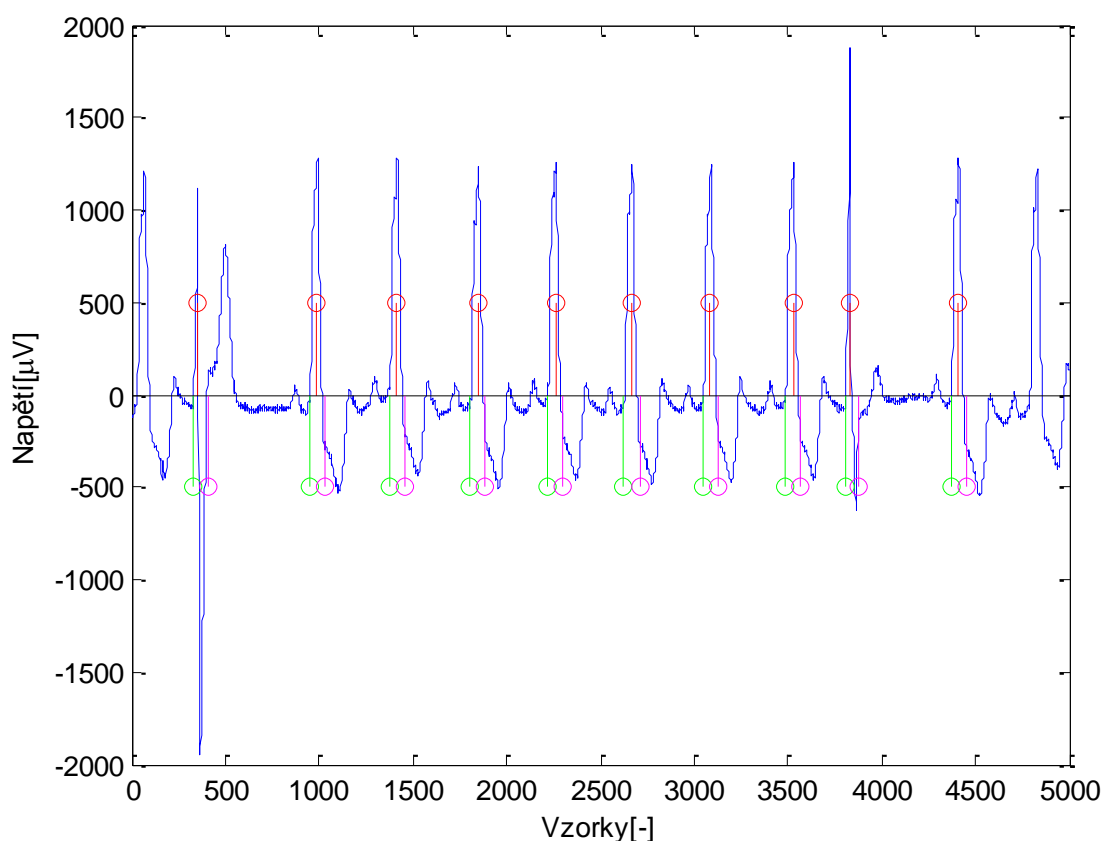


Obrázek 15 - Shluková analýza QRS komplexů v signálu w023.mat – výsledný dendrogram

Takovémuto vyhodnocení se v případě ploch, jakožto atributů, nedá ubránit. Plochy QRS komplexů nejsou tedy v tomto případě vhodným atributem pro shlukovou analýzu.

Signál: W065.mat

V celém souboru analyzovaných signálů se vyskytuje několik takových, které ani přes mnoho způsobů nastavení a zvolení mnoha typů atributů nelze vyhodnotit jednoznačně, bez přidaného komentáře. Jedním z takových signálů je signál w065.mat. Ten obsahuje dva patologické QRS komplexy, každý je ovšem jiného typu, na pozicích 1 a 9 (obr. 16). V tomto případě analýza využívající plochy selže, a to z důvodu vysokého rozdílu hodnoty plochy mezi jednotlivými patologickými komplexy (obr. 17). Oproti tomu shluková analýza provedená pomocí atributu tvarového faktoru a R-R intervalů určí výsledek správně, a to tak, že jsou oba patologické komplexy přiřazeny do stejného shluku (obr. 18).



Obrázek 16 - Detekce QRS komplexů v signálu w065.mat

V následujícím obrázku (obr. 17) je použito atributů ploch a R-R intervalů. Je docíleno toho, že jsou oba komplexy přiřazeny do různých shluků a zároveň selže nastavený řez dendrogramem a tím pádem i barevné oddělení jednotlivých shluků. To je způsobeno v důsledku rozdílné plochy obou QRS komplexů (tab. 11).

Tabulka 11 - Hodnoty atributů ploch a R-R intervalů pro signál w065.mat

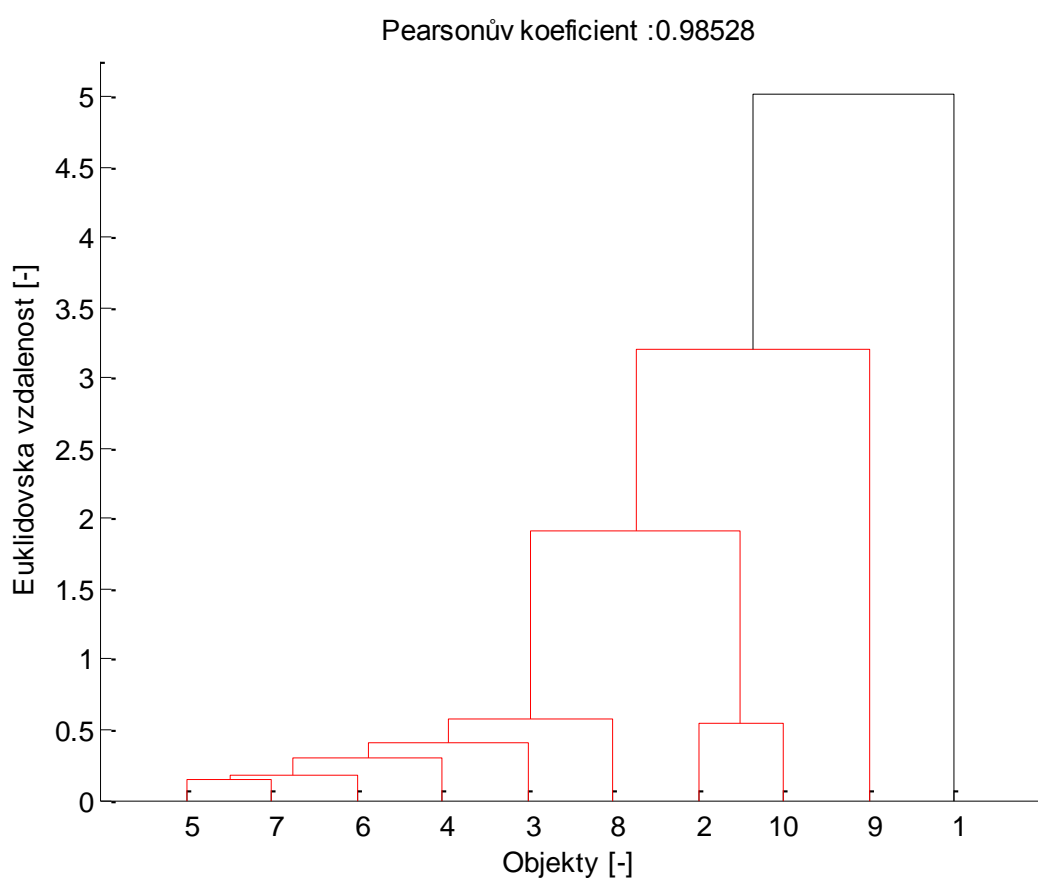
QRS komplex	1	2	3	4	5	6	7	8	9	10
Plochy II	61215	49247	50866	48327	49030	48203	48525	48457	33121	49574
Plochy V2	59001	48312	45364	43857	44970	44529	45185	42257	43046	47683
Plochy V6	52471	9575,4	8246	9276,5	9321,8	8687,5	7921,5	8190,2	26743	9398,5
RR interval	285	638	426	432	412	408	420	438,5	302,5	581,5

Oproti tomu hodnoty tvarových faktorů obou komplexů jsou si díky výpočtu první a druhé difference QRS komplexů podobné. To má za následek přiřazení obou komplexů do jednoho shluku. Tomuto napomáhají i podobné hodnoty R-R intervalů (tab. 12). Dendrogram vytvořený za pomoci shlukové analýzy atributů tvarových faktorů a R-R intervalů je na obrázku dole (obr. 18).

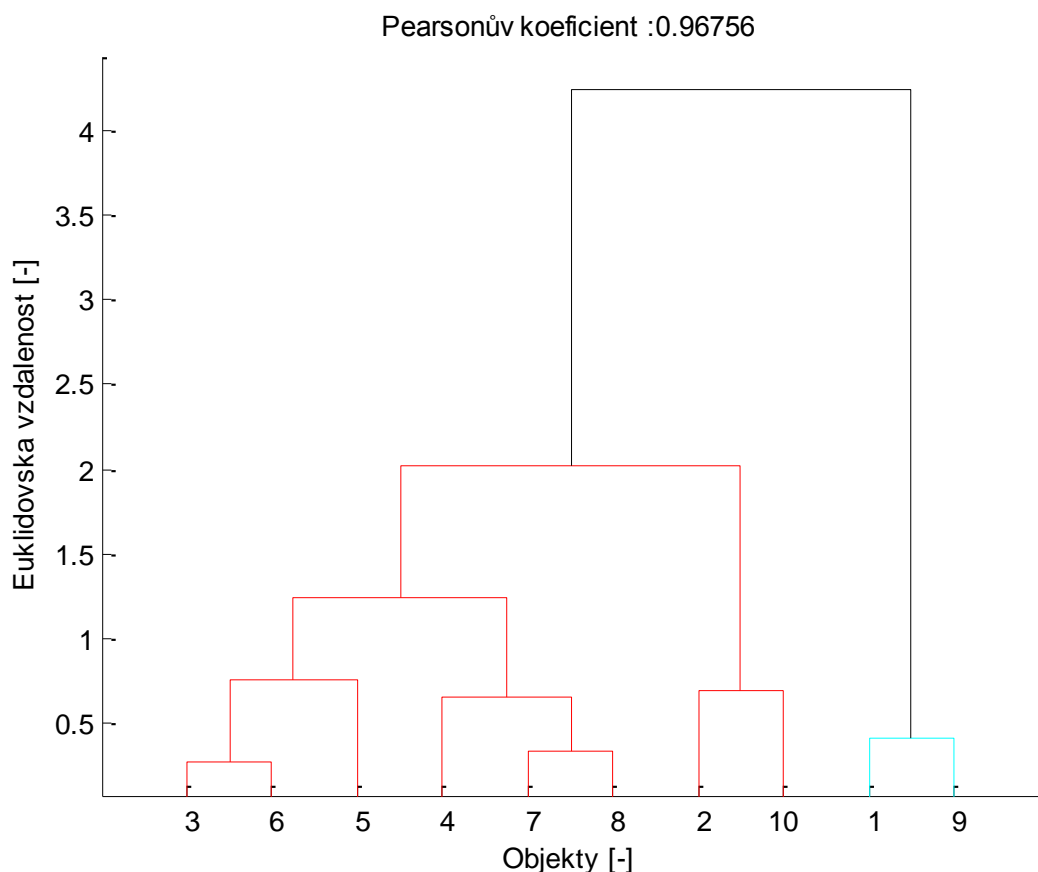
V tomto případě je tedy jedinou správnou možností analýza využívající tvarového faktoru jakožto atributu (obr. 18).

Tabulka 12 - Hodnoty atributů tvarových faktorů a R-R intervalů pro signál w065.mat

QRS komplex	1	2	3	4	5	6	7	8	9	10
FF II	2,741	4,7328	5,2692	5,4598	5,8087	5,4385	5,0423	5,3484	2,7083	5,1901
FF V2	2,3523	5,1288	5,5019	4,6147	5,4397	5,3343	5,355	5,2749	2,4322	4,9188
FF V6	2,7294	4,8268	4,5752	3,7851	5,104	4,6264	3,9346	3,8891	3,0242	4,7899
RR interval	285	638	426	432	412	408	420	438,5	302,5	581,5



Obrázek 17 - Dendrogram pro signál w065.mat s primárním atributem ploch QRS komplexů



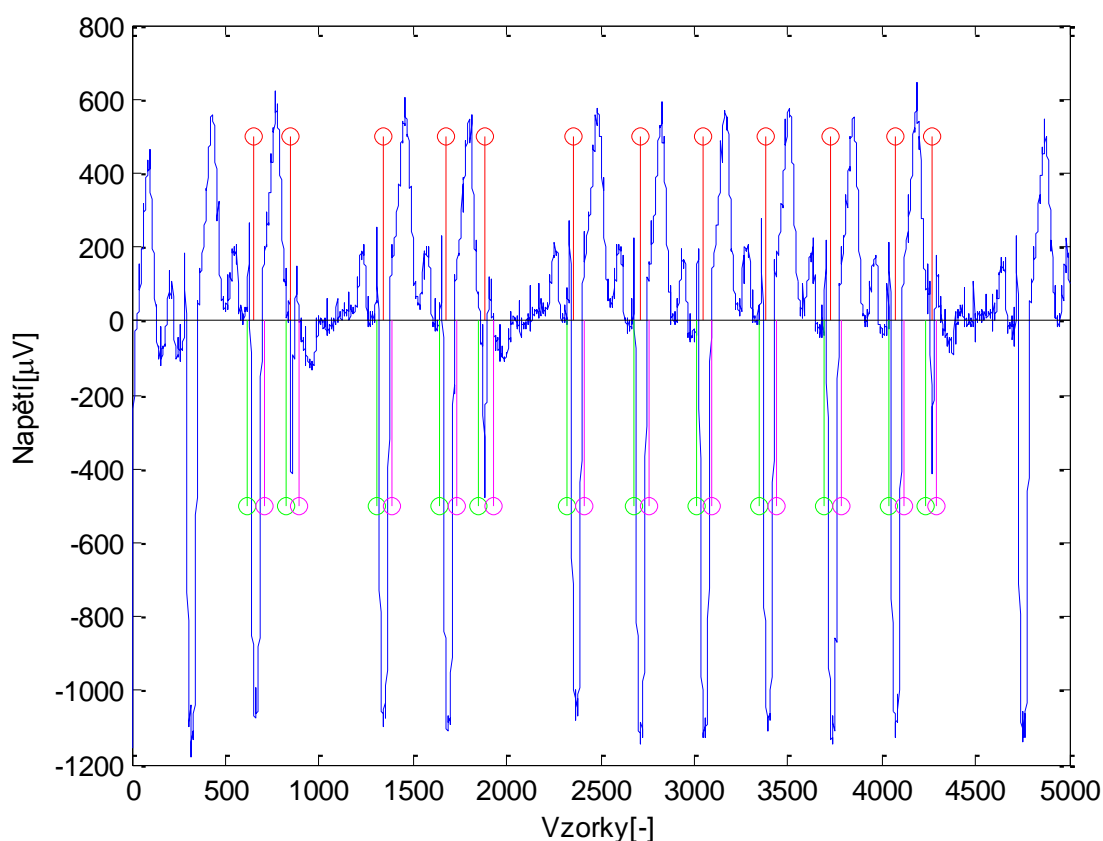
Obrázek 18 - Dendrogram pro signál w065.mat s primárním atributem tvarového faktoru QRS komplexů

6.2 Použití atributu R-R intervalů a tvarového faktoru

6.2.1 Ukázky úspěšného shlukování

Signál: W105.mat

Pro ukázkou správné funkčnosti byl zvolen signál w105.mat, obsahující abnormální komplexy na pozicích 2, 5 a 12 (obr. 19). V tomto případě se jedná o komplexy extrasystolické, u kterých je velké riziko, že vlna R padne do vlny T, což by mohlo vést ke vzniku fibrilace. U použití tvarového faktoru je předpoklad správného přiřazení do patologického shluku hlavně díky odlišnému tvaru extrasystol od fyziologických komplexů.

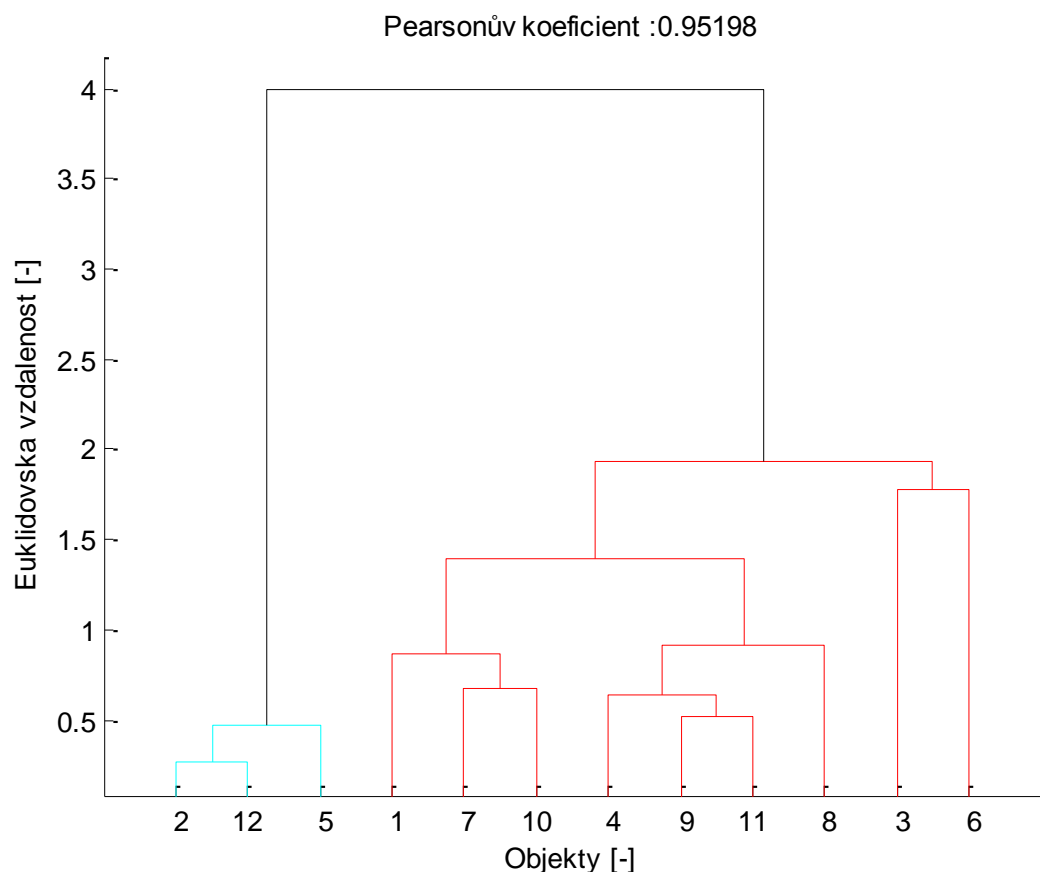


Obrázek 19 - Detekce jednotlivých vln QRS komplexů v signálu w105.mat

Jak je již zmíněno výše, tvar křivek abnormálních komplexů se velmi liší od tvaru komplexů fyziologických, jak je ze signálu na první pohled patrné (obr. 19). Zároveň u patologických QRS komplexů chybí vlna P a patologický cyklus nastupuje mnohem rychleji, než cykly fyziologické, což vede ke zvýraznění důležitosti R-R intervalu jakožto atributu. O tom je možno se přesvědčit v následující tabulce (tab. 13). Hodnoty tvarového faktoru jsou o mnoho menší u komplexů patologických oproti komplexům fyziologickým, což implikuje fakt, že jsou tyto patologické komplexy od komplexů fyziologických velmi tvarově odlišné.

Tabulka 13 - Hodnoty tvarových faktorů a RR intervalů pro signál w105.mat

QRS komplex	1	2	3	4	5	6	7	8	9	10	11	12
FF II	5,1016	4,5070	6,5870	6,7692	4,8170	4,9889	5,5206	6,8428	6,4150	5,9005	6,5035	4,4757
FF V2	6,4566	3,1671	6,3023	6,2170	3,1709	6,5325	5,8121	6,7793	5,6897	6,0339	6,3500	2,7841
FF V6	6,6976	3,8689	6,5961	6,3356	3,4702	6,4202	5,9894	7,7183	6,7927	6,7197	7,1573	3,7855
RR interval	339,5	198,5	487,5	341	200,5	481	348	336	342	341	342,5	195



Obrázek 20 - Dendrogram vytvořený za pomoci tvarových faktorů a R-R intervalů pro signál w105.mat

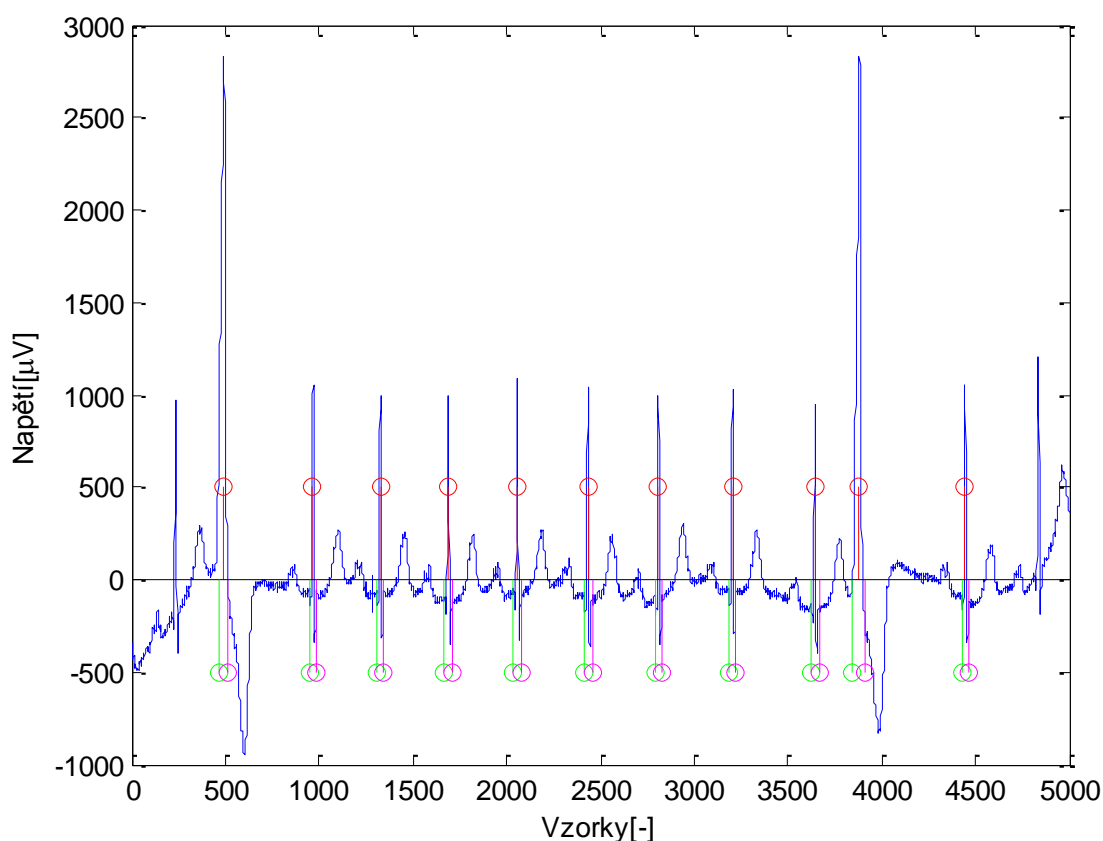
Komplex č. 5 má díky vyšší hodnotě tvarových faktorů větší distanci ke shluku patologických komplexů č. 2 a 12 (obr. 20).

Signál: W061.mat

Dalším ze signálů, u kterých proběhne shlukování úspěšně, je signál w061.mat. Tento signál obsahuje dvě velmi zřetelné patologie na pozicích č. 1 a č. 10 (obr. 21). Shluková analýza je v tomto případě provedena pomocí shlukovací metody UPGMA.

Tabulka 14 - Tabulka hodnot pro kombinaci atributů tvarového faktoru a R-R intervalu pro signál w061.mat

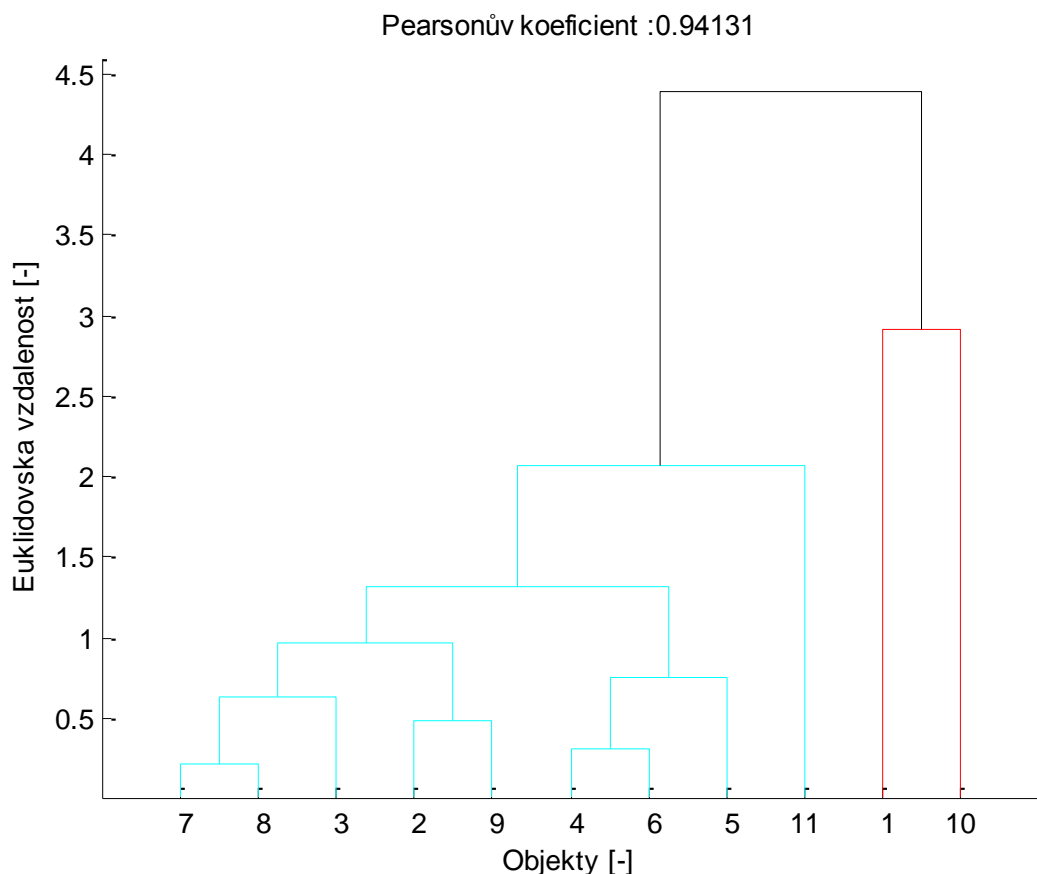
QRS komplex	1	2	3	4	5	6	7	8	9	10	11
FF II	1,8886	1,5286	1,4707	1,5379	1,5327	1,6058	1,5195	1,5235	1,5367	2,402	1,5678
FF V2	3,6837	1,5563	1,4612	1,4696	1,4436	1,447	1,5033	1,4972	1,4933	3,1399	1,4801
FF V6	2,1051	1,8192	1,7323	2,0653	2,2367	2,0405	1,8495	1,8759	1,888	2,6359	2,0983
RR interval	256	476	358	363	368	376	377	394	441	238	562



Obrázek 21 - Detekce vln jednotlivých komplexů v signálu w061.mat

Hodnoty tvarových faktorů v tomto případě hrají ve shlukové analýze roli srovnatelnou s hodnotami R-R intervalů. U prvního a desátého komplexu dojde ke zkrácení R-R intervalu oproti fyziologickým cyklům a jejich R-R intervalům, tento atribut je tedy pro shlukování směrodatný a stejně tak u tvarových faktorů obou cyklů dojde ke vzniku podstatného rozdílu v hodnotách oproti cyklům fyziologickým (tab. 14).

U této analýzy je tedy možno již při pohledu na samotný signál (obr. 21) a na optický rozdíl ve tvaru a velikosti vln č. 1 a 10. předpokládat, že obě vlny budou za pomoci atributu hodnot tvarového faktoru přiřazeny do samostatného shluku. Tuto teorii dokazuje výsledný dendrogram (obr. 22), jenž skutečně zobrazuje přiřazení obou cyklů do samostatného shluku. Rovněž hodnota euklidovské vzdálenosti mezi shluky normálními a abnormálními je dle oddělení shluků v dendrogramu vysoká.



Obrázek 22 - Výsledný dendrogram pro signál w061.mat

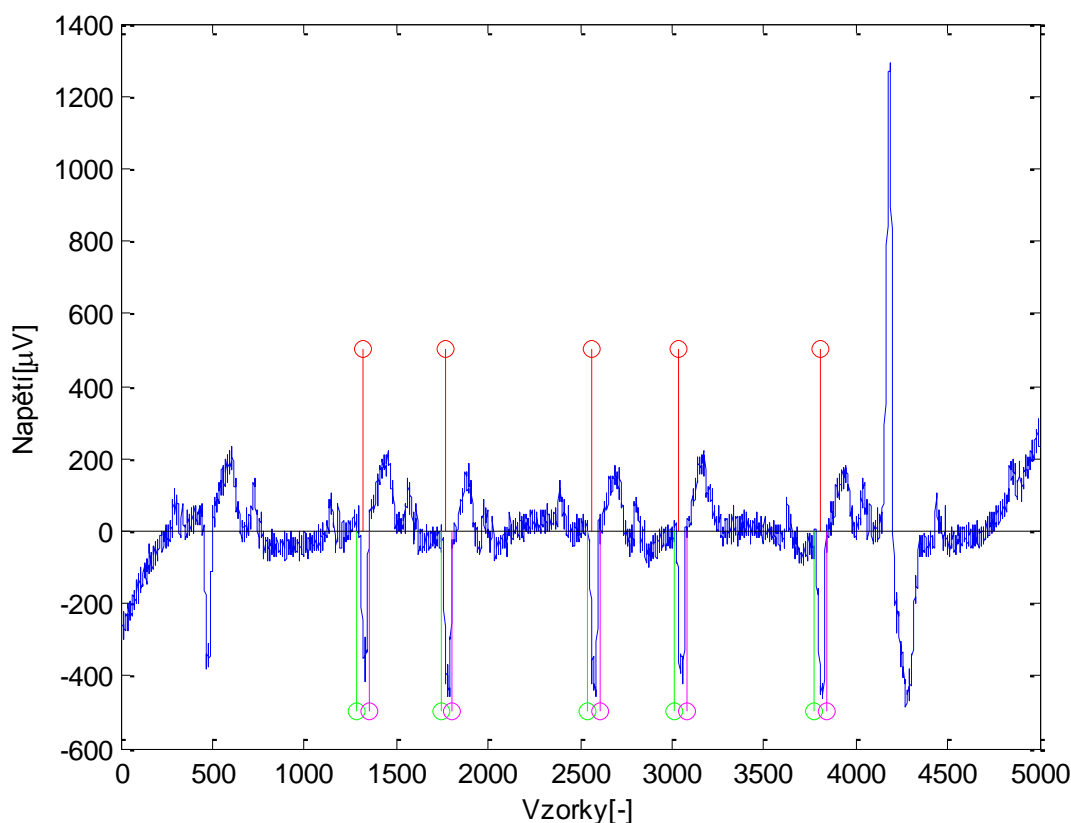
6.2.2 Ukázky neúspěšného shlukování

Signál: W054.mat

Špatná funkce metody shlukování, za použití tvarového faktoru jakožto atributu, může být prezentována signálem w054.mat ze sbírky signálů. Tento signál obsahuje abnormální komplexy na detekovaných pozicích 2 a 4, a jednu okrajovou nedetekovanou patologii na pozici 6 vzhledem k detekci (obr. 23). U tohoto signálu se vyskytuje i možnost, že i přes rozdílnost délky R-R intervalů u komplexů 2 a 4 a zvláštnosti tvaru těchto komplexů, se ve skutečnosti nejedná o patologické komplexy. Je tedy možné, že ač komplexy 2 a 4 přicházejí velmi brzy po předchozích QRS komplexech, jedná se pouze o formu zrychlení srdeční činnosti, a jediná vyskytující se patologie je na pozici č. 6.

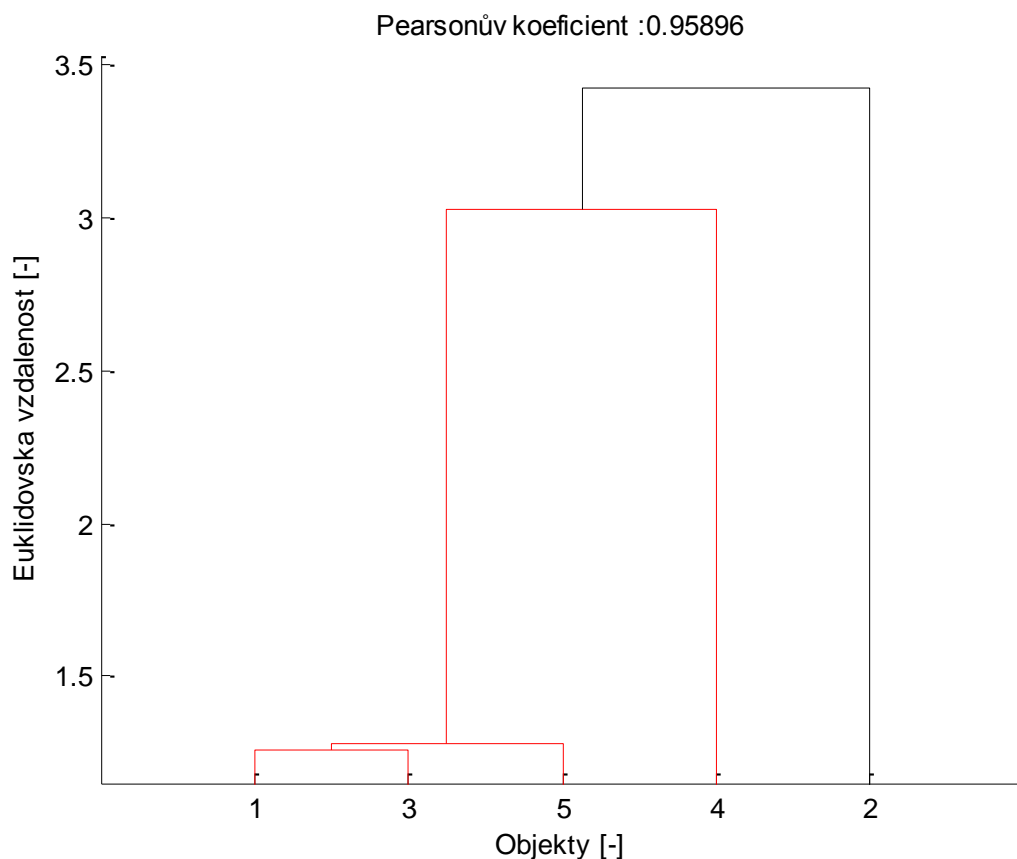
V tomto vyhodnocení však předpokládáme, že se na pozicích 2 a 4 skutečně nacházejí vlny patologické. Hodnoty tvarových faktorů se od sebe mezi pozicemi 2 a 4 mírně liší, což i

přes podobnost v krátké vzdálenosti R-R intervalu způsobí, že je každá z abnormalit přiřazena do samostatného shluku (obr. 24). R-R interval tedy v tomto případě není dostačujícím atributem k vyhodnocení podobnosti mezi těmito vlnami. Podobná chyba, kdy není detekován jeden z patologických komplexů, případně jediný, který se v signálu vyskytuje, je přítomna například u signálu w120.mat.



Obrázek 23 - Detekce jednotlivých vlny komplexů v signálu w054.mat

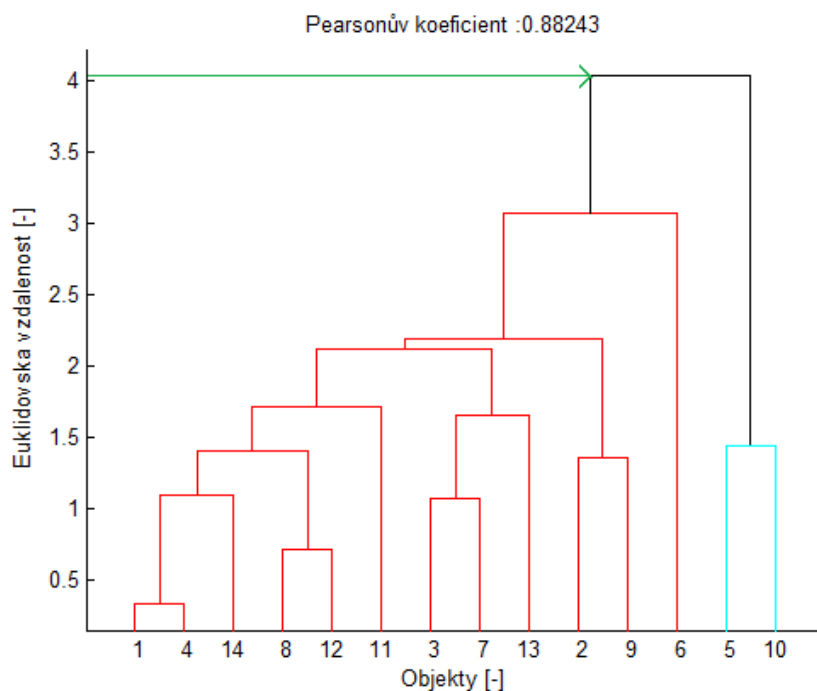
U výsledného dendrogramu je řez proveden tak, že všechny fyziologické komplexy jsou přiřazeny do samostatného shluku společně s patologickým komplexem č. 4. a další potencionálně patologický cyklus (č. 2) je přiřazen do samostatného shluku (obr. 24). Došlo-li by k detekci posledního komplexu v tomto signálu, dá se na základě výsledku předchozích analýz předpokládat, že by byl tento komplex na základě odlišnosti v tvarovém faktoru přiřazen do samostatného shluku.



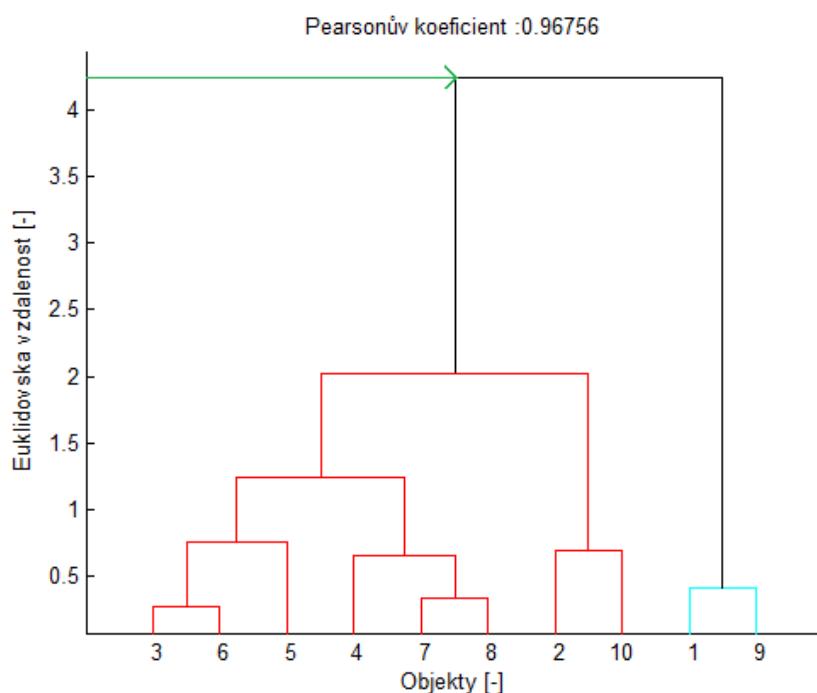
Obrázek 24 - Výsledný dendrogram pro signál w054.mat

6.3 Poznámka k nastavení prahu řezu dendrogramem a ke speciálnímu způsobu rozměření některých signálů.

Díky množství signálů, a z nich vzešlých dendrogramů, bylo zapotřebí nějakým způsobem získat hodnotu prahu pro řez těmito dendrogramy. Díky standardizaci dat a použití funkce `pdist` pro výpočet Euklidovských vzdáleností mezi objekty je dosaženo určitých hodnot, na jejichž základě je možno stanovit práh řezu tak, aby byla analýza skutečně co nejvíce automatická. Nejvhodnější přibližnou hodnotou prahu pro řez je hodnota maximální distance mezi posledními dvěma shluky v analýze. Jedná se o nejvhodnější způsob stanovení prahu řezu a to z toho důvodu, že téměř vždy maximální hodnota euklidovské distance mezi uzly značí vzdálenost oddělující od sebe shluky patologických a fyziologických cyklů v analýze. Díky tomuto faktu je následně dendrogram rozdělen většinou na dva barevně oddělené shluky. Výhodou této metody určení prahu řezu je fakt, že maximální hodnota pro práh řezu je vždy k dispozici, nehledě na velikost dosažených distancí či na změnu metody. Tento princip nastavení prahu je využit jak u jednosvodových analýz, tak i u analýz pocházejících ze tří pseudoortogonálních svodů.

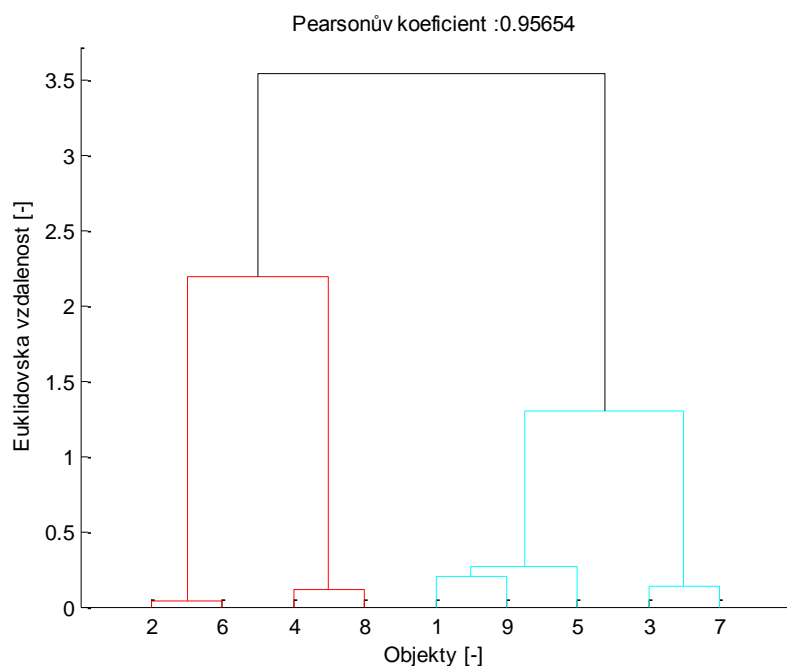


Obrázek 25 - Určení prahu řezu pro signál w047.mat

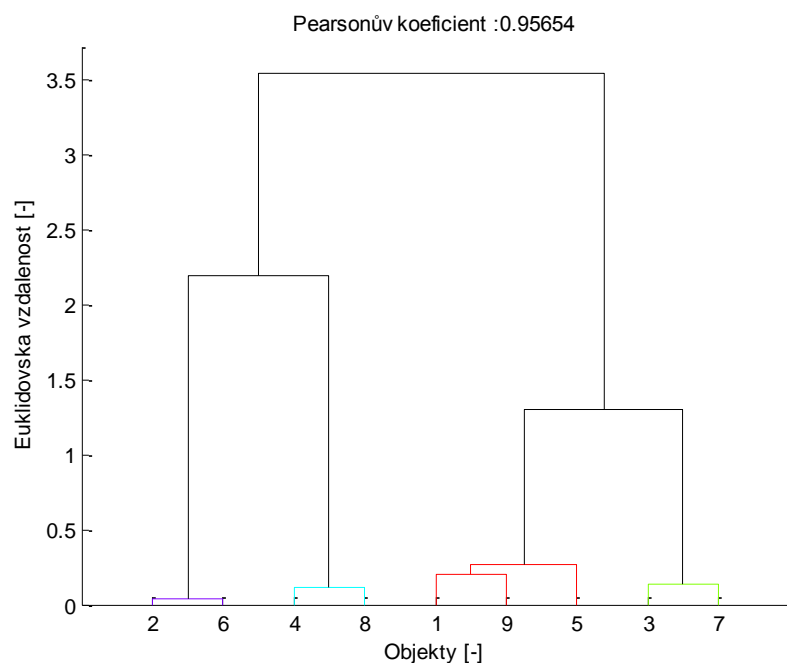


Obrázek 26 – Určení prahu řezu pro signál w065.mat

Na obrázcích (obr. 25) a (obr. 26) jsou vybrané prahy pro řez vyznačeny zeleně. U signálu w117.mat je však například pro správně rozřazení do shluků nutno zvolit jiný práh řezu dendrogramem, a to ne maximální, ale průměrnou hodnotu euklidovských distancí mezi shluky. Tento princip je aplikován pro demonstraci vlivu řezu na výsledek shlukové analýzy (obr. 28).



Obrázek 27 - Výsledný dendrogram pro signál w117.mat s použitím klasického prahu pro řez



Obrázek 28 - Výsledný dendrogram pro signál w117.mat po manuálním nastavení prahu

Další speciální případ vyhodnocení je uzpůsoben fakt, že díky detekci, nebo naopak absenci detekce některých vln nemohlo dojít ke správnému rozměření a tedy vyhodnocení signálu. Proto je zapotřebí v případě signálu w034.mat pro obě kombinace atributů použít rozměření ne ze všech dvanácti svodů signálu, ale jen ze svodu č. 4. Všechny ostatní svody totiž buď neobsahují požadované vlny, nebo jich naopak obsahují mnoho, rozměření není úspěšné, vektory vln jsou nekonzistentní, a analýza by v tomto případě nemohla díky chybě

programu proběhnout. Podobně je tomu u signálu w117.mat, kdy u posledního detekovaného komplexu (tedy komplexu č. 10) nedojde k detekci S vlny, a tedy ke správnému rozměření. Hodnoty ploch a tvarových faktorů potom tedy u tohoto komplexu nelze brát jako směrodatné, a analýza se tedy jeví jako neúspěšná. Proto je nutné pro správnost shlukování tuto vlnu z analýzy odstranit. Tomuto účelu slouží samostatná funkce shlukovani_w117.m.

K poslednímu odlišnému způsobu vyhodnocování shlukování vede fakt, že u některých signálů, označených v tabulce (tab. 6 a 7) fialově, je zapotřebí použití jiné metody shlukování, než je obvyklá metoda UPGMA. Tato matematická metoda u daných signálů nevede k přiřazení správných objektů do příslušných shluků, je tedy potřeba použít jinou matematickou metodu. Ve většině případů se jedná o Wardovu shlukovací metodu.

7. Závěr

Hlavním cílem této bakalářské práce byl výběr vhodných příznaků (atributů) a následné provedení shlukové analýzy u série signálu, obsahujících patologické komplexy vln. Byla tedy vytvořena série atributů, pomocí jejichž kombinací bylo shlukování prováděno. Tyto analýzy byly zprvu prováděny pouze z prvního svodu signálu (tab. 6), až později byla analýza rozšířena o další dva svody. Použití tří pseudoortogonálních svodů (svody II, V2 a V6) se v konečném důsledku ukázalo jako vhodné a byla tedy vytvořena tabulka, porovnávající výsledky analýzy pomocí dvou kombinací atributů pro tyto svody (tab. 7).

Pro první shlukovou analýzu bylo použito atributů R-R intervalů a tvarového faktoru, kdy tato kombinace atributů vycházela z diplomové práce Ing. Chrobáka. [7] Shluková analýza ze tří svodů byla v tomto případě úspěšná, a přibližně odpovídající výsledkům, jež byly uvedeny ve výše zmíněné diplomové práci. Rozdíl ve výsledcích mohl vzniknout díky faktu, že Ing. Chrobák používal R-R interval následující po cyklu (tedy vliv kompenzačních pauz), zatímco v této práci jsou použity intervaly před cykly (rychlejší nástup extrasystolických cyklů), nebo mohlo odchylení se od předchozích výsledků analýzy vzniknout v důsledku faktu, že v této práci se pro výpočet tvarových faktorů využívalo signálu bez filtrace. Stejně tak bylo u několika signálů zapotřebí změny způsobu rozměření, protože z některých svodů (případně ze všech) nebylo rozměření provedeno korektně a docházelo tedy k chybám vlivem nekonzistence vektorů v matici v prostředí MATLAB. Všechny výše zmíněné případy jsou popsány v podkapitole 6.3, která se touto problematikou zabývá.

Pro druhý případ shlukové analýzy bylo použito ploch QRS komplexů ze tří pseudoortogonálních svodů signálu EKG a R-R intervalu z těchto svodů. Plocha těchto komplexů byla získána převedením hodnot napětí v intervalu Q a S do hodnot absolutních a následnou integrací těchto hodnot. Plochy QRS komplexů byly vybrány z toho důvodu, že ve většině případů vede patologie v EKG signálu buď ke snížení nebo naopak zvýšení napěťového rozkmitu signálu a tedy ke zmenšení či zvětšení hodnoty plochy celého komplexu. Analýza, provedená za pomoci těchto atributů podala o něco lepší výsledky, než v prvním případě (tedy za pomoci tvarového faktoru a R-R intervalů), avšak taktéž obsahuje některé nedostatky, týkající se několika málo signálů. Stejně jako v prvním případě, i zde bylo potřeba u několika signálů použít specifický způsob rozměření, případně vynechání chybné vlny (např. signál w117.mat), popřípadě bylo zapotřebí změnit způsob metody shlukování či práh řezu dendrogramem. Samotným problémem atributu ploch QRS komplexů však je, jak již je popsáno výše v kapitole zabývající se výběrem atributů, že plochy patologických komplexů se například díky rušení či díky jinému typu patologie v signálu mohou lišit a v důsledku tohoto dochází buď ke špatnému přiřazení komplexů do shluku, či přiřazení dvou patologických komplexů do dvou samostatných shluků. Pak je potřeba využít některý ze

způsobů speciálního vyhodnocení či vytvoření dendrogramu, zmíněných v podkapitole 6.3. Jak je vidno z tabulek (tab. 6) a (tab. 7), vedlo použití třísvodové shlukové analýzy k potlačení potřeby měnit u mnoha signálů metodu shlukování a přispělo k lepším výsledkům shlukové analýzy.

Některé z výsledků shlukování jsou popsány v kapitolách 6.1 a 6.2 a v příslušných podkapitolách a další výsledky jsou uvedeny v příloze této práce, jelikož díky množství signálů, použitých jako vstup pro tuto analýzu, není možné popsat obsáhleji výsledek pro každý z nich.

Ne ve všech případech však byla analýza za pomoci zvolených příznaků (atributů) úspěšná pro všechny použité EKG signály v souboru signálů CSE, což je způsobeno nemožností obsáhnout všechny typy patologií do zvolených atributů, s jejichž pomocí je analýza prováděna. Pro dané téma této práce tedy zajisté existuje celá řada dalších možností, jak se dá shluková analýza EKG signálu dále zlepšit a které nejsou popsány díky restrikci rozsahu práce. Množství možností analýzy je amplifikováno faktem, že signál obsahuje dvanáct různých svodů, ze kterých je možno získávat popisné parametry. Za povšimnutí stojí fakt, že ve většině signálů se patologie projevuje změnou tvaru či pozice vlny T, případně její absolutní absencí. Bylo by také vhodné se v této problematice dále zaměřit na rozkmit mezi vlnami S a T v komplexu, který se u patologických komplexů poměrně liší od fyziologických.

8. Použitá literatura

- [1] MOUREK, J. *Fyziologie -- učebnice pro studenty zdravotnických oborů*. 1. vyd. Grada, 2005. ISBN 978-80-247-1190-4.
- [2] KAŇKOVÁ, K. *Patologická fyziologie pro bakalářské studijní programy*. 2. vyd. Brno: Masarykova univerzita, 2009, 164 s. ISBN 978-80-210-4923-9.
- [3] HAMPTON, J. R. *EKG stručně, jasně, přehledně*. 2., rozš. vyd. Praha: Grada, 2005, 149 s. ISBN 80-247-0960-0.
- [4] LUKL, J. *Srdeční arytmie v kazuistikách: postupy podle léčebných standardů*. 1. vyd. Praha: Grada, 2006, 116 s. ISBN 80-247-1544-9.
- [5] FARSKÝ, Š. *EKG do vrecka*. Martin: Osveta, 2010, 103 s. ISBN 978-808-0631-789.
- [6] KOZUMPLÍK, J. *Analýza biologických signálů*, (Rozpracovaná skripta), VUT v Brně, 2011.
- [7] CHROBÁK, M. *Shluková analýza*. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2012. 71 s. Vedoucí diplomové práce doc. Ing. Jiří Kozumplík, CSc..
- [8] KAUFMAN, L. a P. J ROUSSEEUW. *Finding groups in data: an introduction to cluster analysis*. Hoboken, N.J.: Wiley, c2005, xiv, 342 p. ISBN 04-717-3578-7.
- [9] ROMESBURG, H. *Cluster analysis for researchers*. Morrisville: Lulu Press, 2004, xii, 334 s. ISBN 14-116-0617-5.
- [10] ŽÁK, L. *Shluková analýza I. Automatizace*. 2004, roč. 47, č. 3, s. 3. Dostupné z: http://www.volny.cz/elzet/Libor/Aut_cl_1.pdf
- [11] RANGAYYAN, R. M. *Biomedical signal analysis: a case-study approach*. New York, N.Y.: Wiley-Interscience, c2002, xxxv, 516 p. ISBN 04-712-0811-6.
- [12] MUŽÍK, J. *Elektromagnetické pole v buňkách: Akční potenciál*. [online]. [cit. 2012-12-03]. Dostupné z: <http://fernet.kbx.cz/other/fel/bup/bup.html>

- [13] Studium.yarousch - studijní materiály pro budoucí zdravotní sestřičky a nejen pro ně: vnitřní lékařství - poruchy srdečního rytmu. [online]. [cit. 2012-12-03]. Dostupné z: http://www.yarousch.cz/studium/body.php?menu=menu_vnitri_lekarstvi&body=vnitri_lekarstvi/vnitri_lekarstvi_004
- [14] KOZUMPLÍK, J. Umělá inteligence v medicíně: Shluková analýza. [online]. [cit. 2012-12-03]. Dostupné z: https://www.vutbr.cz/elearning/file.php/123626/prednasky/Prednasky_12_13/AUIN_02_shlukova_analyza.pdf
- [15] VÍTEK, M. *Automatické rozměření signálů EKG*. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2010. 129 s. Vedoucí dizertační práce doc. Ing. Jiří Kozumplík, CSc. Dostupné z: https://www.vutbr.cz/www_base/zav_prace_soubor_verejne.php?file_id=34194
- [16] KUNA, Z. *Detekce komplexů QRS v signálech EKG*. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2010. 68s. Vedoucí diplomové práce doc. Ing. Jiří Kozumplík, CSc.

Seznam obrázků

Obrázek 1 - Převodní systém srdeční [13]	9
Obrázek 2 - Vznik akčního potenciálu v buňce myokardu [12].....	10
Obrázek 3 - Vlevo: Einthovenovy bipolární končetinové svody I, II a III; vpravo: unipolární Goldbergerovy svody avR, avF a avL[6]	11
Obrázek 4 - Umístění hrudních svodů [6]	11
Obrázek 5 - Základní vlny a intervaly v EKG [6]	13
Obrázek 6 - Rozdělení shlukové analýzy [14]	15
Obrázek 7 - Výsledný dendrogram [9].....	20
Obrázek 8 - Vytvoření všech možných shluků a výpočet jejich hodnot E [9].....	22
Obrázek 9 - Ukázka rozdílnosti délek R-R intervalu u patologické a fyziologické vlny.....	26
Obrázek 10 - Detekce QRS vln v signálu w075.mat.....	28
Obrázek 11 - Provedení shlukové analýzy pro signál w075.mat metodou UPGMA.....	29
Obrázek 12 - Ukázka detekce QRS vln.....	30
Obrázek 13 - Samotné shlukování pro signál W122.mat pomocí atributů R-R intervalu a ploch QRS komplexů	31
Obrázek 14 - Detekce QRS komplexů v signálu w023.mat.....	32
Obrázek 15 - Shluková analýza QRS komplexů v signálu w023.mat – výsledný dendrogram.....	33
Obrázek 16 - Detekce QRS komplexů v signálu w065.mat.....	34
Obrázek 17 - Dendrogram pro signál w065.mat s primárním atributem ploch QRS komplexů	35
Obrázek 18 - Dendrogram pro signál w065.mat s primárním atributem tvarového faktoru QRS komplexů.....	36
Obrázek 19 - Detekce jednotlivých vln QRS komplexů v signálu w105.mat.....	37
Obrázek 20 - Dendrogram vytvořený za pomoci tvarových faktorů a R-R intervalů pro signál w105.mat	38
Obrázek 21 - Detekce vln jednotlivých komplexů v signálu w061.mat	39
Obrázek 22 - Výsledný dendrogram pro signál w061.mat.....	40
Obrázek 23 - Detekce jednotlivých vln komplexů v signálu w054.mat	41
Obrázek 24 - Výsledný dendrogram pro signál w054.mat.....	42
Obrázek 25 - Určení prahu řezu pro signál w047.mat	43
Obrázek 26 – Určení prahu řezu pro signál w065.mat.....	43
Obrázek 27 - Výsledný dendrogram pro signál w117.mat s použitím klasického prahu pro řez	44
Obrázek 28 - Výsledný dendrogram pro signál w117.mat po manuálním nastavení prahu	44
Obrázek 29 - Detekce vln v signálu W006.mat.....	55

Obrázek 30 - Neúspěšná shluková analýza signálu w006.mat (kombinace atributu plochy a R-R intervalu).....	55
Obrázek 31 - Neúspěšná shluková analýza signálu w006.mat (kombinace atributu tvarového faktoru a R-R intervalu)	56
Obrázek 32 - Detekce vln v signálu w054.mat	56
Obrázek 33 - Neúspěšná shluková analýza signálu w054.mat (kombinace atributu plochy a R-R intervalu).....	57
Obrázek 34 - Detekce vln v signálu w120.mat	57
Obrázek 35 - Neúspěšná shluková analýza signálu w120.mat (kombinace atributu plochy a R-R intervalu).....	58
Obrázek 36 - Neúspěšná shluková analýza signálu w120.mat (kombinace atributu tvarového faktoru a R-R intervalu)	58
Obrázek 37 - Neúspěšná shluková analýza signálu w028.mat (kombinace atributu tvarového faktoru a R-R intervalu)	59
Obrázek 38 - Detekce vln v signálu w103.mat	59
Obrázek 39 - Neúspěšná shluková analýza signálu w103.mat (kombinace atributu plochy a R-R intervalu).....	60
Obrázek 40 - Neúspěšná shluková analýza signálu w103.mat (kombinace atributu tvarového faktoru a R-R intervalu)	60
Obrázek 41 - Detekce vln v signálu w047.mat	61
Obrázek 42 - Úspěšná shluková analýza signálu w047.mat (kombinace atributu ploch a R-R intervalu)	61
Obrázek 43 - Úspěšná shluková analýza signálu w047.mat (kombinace atributu tvarového faktoru a R-R intervalu)	62
Obrázek 44 - Detekce vln v signálu w026.mat	62
Obrázek 45 - Úspěšná shluková analýza signálu w026.mat (kombinace atributu ploch a R-R intervalu)	63
Obrázek 46 - Úspěšná shluková analýza signálu w026.mat (kombinace atributu tvarového faktoru a R-R intervalu)	63
Obrázek 47 - Blokové schéma programu pro provedení shlukové analýzy.....	64

Seznam tabulek

Tabulka 1 - Trvání jednotlivých vln a komplexů EKG křivky a původ jejich vzniku [1]	12
Tabulka 2 - Výběr nejmenší hodnoty podobnosti v matici podobností [9].....	18
Tabulka 3 - Nově získaná matice podobností [9].....	19
Tabulka 4 - Výsledná matice podobností [9]	19
Tabulka 5 - Vstupní data [9]	21

Tabulka 6 - Výsledky jednosvodové analýzy z prvního svodu signálu pro různé kombinace atributů	27
Tabulka 7 - Výsledky třísvodové analýzy z pseudoortogonálních svodů pro různé kombinace atributů	27
Tabulka 8 - Hodnoty ploch a RR intervalů pro signál w075.mat	29
Tabulka 9- Hodnoty ploch a RR intervalů pro signál w122.mat.....	30
Tabulka 10 - Atributy signálu w023.mat	32
Tabulka 11 - Hodnoty atributů ploch a R-R intervalů pro signál w065.mat.....	34
Tabulka 12 - Hodnoty atributů tvarových faktorů a R-R intervalů pro signál w065.mat	35
Tabulka 13 - Hodnoty tvarových faktorů a RR intervalů pro signál w105.mat.....	37
Tabulka 14 - Tabulka hodnot pro kombinaci atributů tvarového faktoru a R-R intervalu pro signál w061.mat	38
Tabulka 15 - Hodnoty Pearsonových korelačních koeficientů pro obě kombinace atributů, modře vyznačené hodnoty značí přílišné zkreslení dendrogramem.....	54

Seznam zkratk, symbolů a příloh

A. Seznam zkratk

VF	<i>Vzdálenostní faktor</i>
FF	<i>Tvarový faktor</i>
ATP	<i>Adenosintrifosfát</i>
EKG	<i>Elektrokardiogram</i>
UPGMA	<i>Unweighted Pair Group Method with Arithmetic Mean</i>
CLINK	<i>Complete Linkage Clustering Method</i>
SLINK	<i>Single Linkage Clustering Method</i>
WARD	<i>Wardova shlukovací metoda</i>

B. Seznam příloh

Příloha č. 1 : Hodnoty Pearsonových korelačních koeficientů	54
Příloha č. 2 : Další ukázky chybné shlukové analýzy pro obě kombinace atributů	55
Příloha č. 3 : Další ukázky úspěšné shlukové analýzy pro obě kombinace atributů.....	61
Příloha č. 4 : Blokové schéma programu pro provedení shlukové analýzy	64
Příloha č. 5 : Přílohy na CD	64

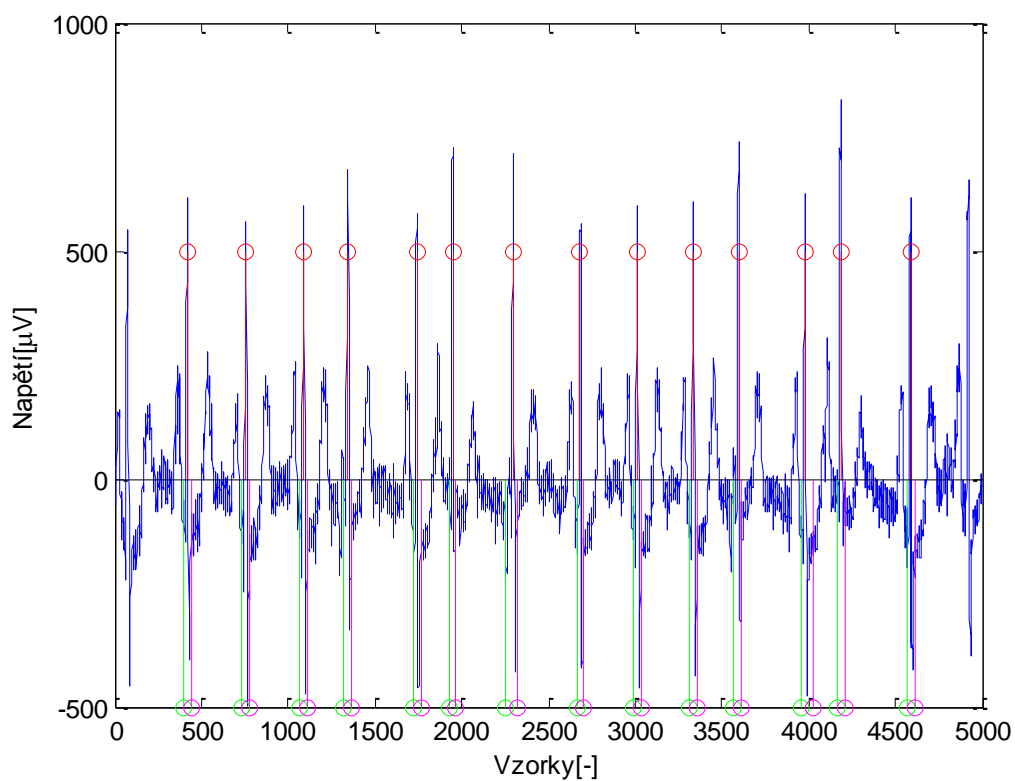
Přílohy

Příloha č. 1 : Hodnoty Pearsonových korelačních koeficientů

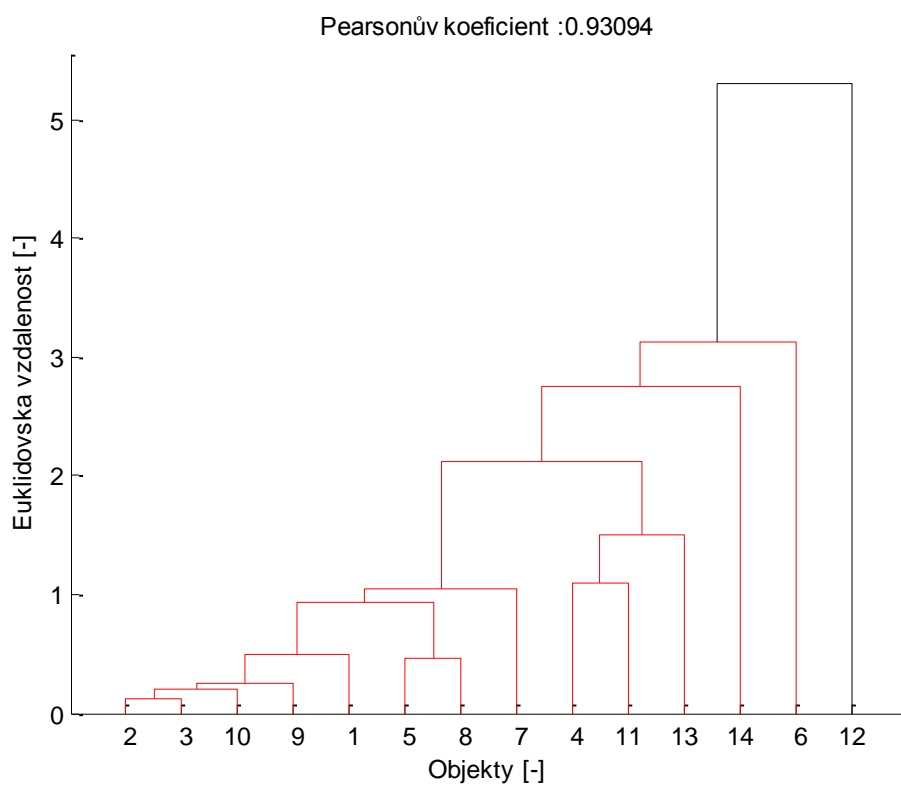
Tabulka 15 - Hodnoty Pearsonových korelačních koeficientů pro obě kombinace atributů, modře vyznačené hodnoty značí přílišné zkreslení dendrogramem

Hodnoty Pearsonových korelačních koeficientů					
Plochy + R-R interval			Tvarový faktor + R-R interval		
Signál	Metoda	Hodnota	Signál	Metoda	Hodnota
w006.mat	UPGMA	0,9309	w006.mat	UPGMA	0,8440
w020.mat	UPGMA	0,7744	w020.mat	UPGMA	0,9178
w023.mat	UPGMA	0,9657	w023.mat	UPGMA	0,7707
w026.mat	UPGMA	0,9907	w026.mat	UPGMA	0,9002
w028.mat	UPGMA	0,9882	w028.mat	UPGMA	0,9560
w034.mat	UPGMA	0,9835	w034.mat	UPGMA	0,9137
w040.mat	UPGMA	0,9875	w040.mat	UPGMA	0,9979
w047.mat	UPGMA	0,8728	w047.mat	UPGMA	0,8824
w054.mat	UPGMA	0,6411	w054.mat	UPGMA	0,9590
w061.mat	UPGMA	0,9743	w061.mat	UPGMA	0,9413
w065.mat	UPGMA	0,9853	w065.mat	UPGMA	0,9676
w074.mat	UPGMA	0,9290	w074.mat	UPGMA	0,9696
w075.mat	UPGMA	0,9811	w075.mat	UPGMA	0,9886
w103.mat	UPGMA	0,9688	w103.mat	UPGMA	0,9708
w105.mat	UPGMA	0,9523	w105.mat	UPGMA	0,9520
w109.mat	WARD	0,8525	w109.mat	UPGMA	0,9015
w112.mat	UPGMA	0,9711	w112.mat	UPGMA	0,9969
w115.mat	UPGMA	0,9886	w115.mat	UPGMA	0,9693
w117.mat	UPGMA	0,9415	w117.mat	UPGMA	0,8601
w120.mat	UPGMA	0,7196	w120.mat	UPGMA	0,7943
w122.mat	UPGMA	0,8868	w122.mat	UPGMA	0,9628

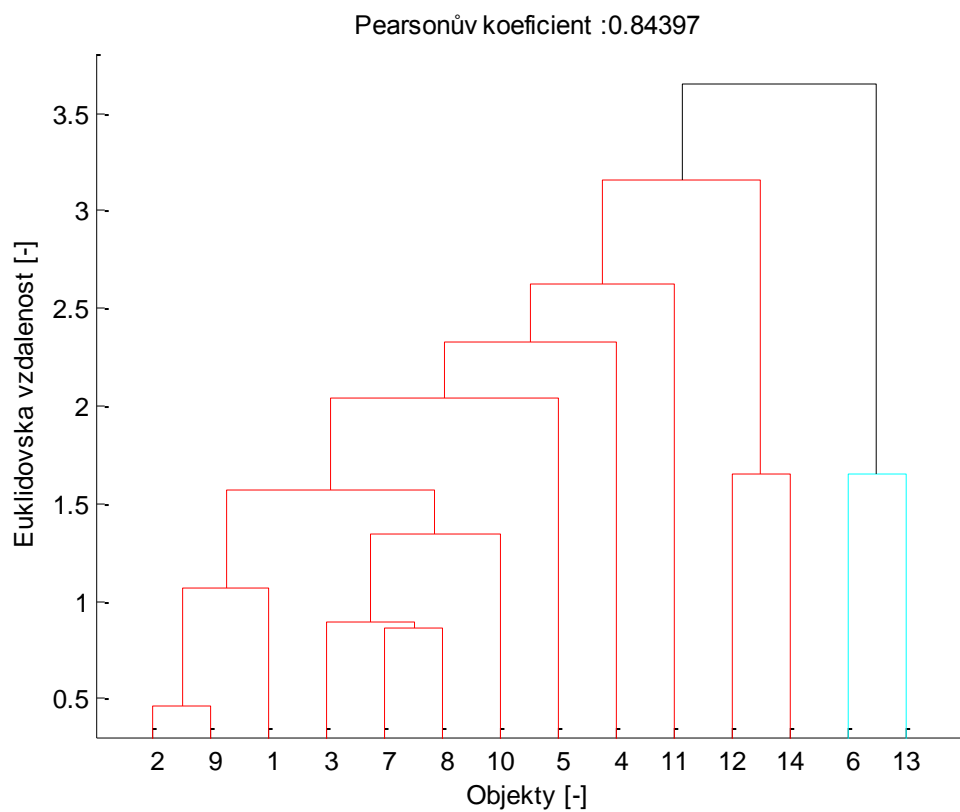
Příloha č. 2 : Další ukázky chybné shlukové analýzy pro obě kombinace atributů



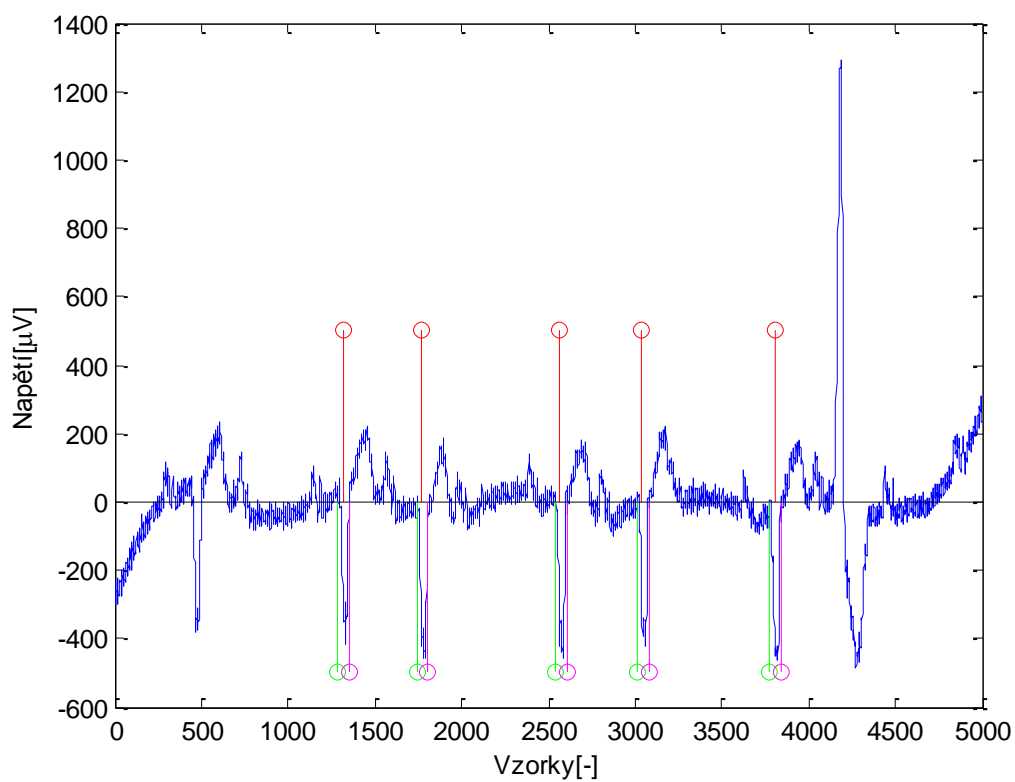
Obrázek 29 - Detekce vln v signálu W006.mat



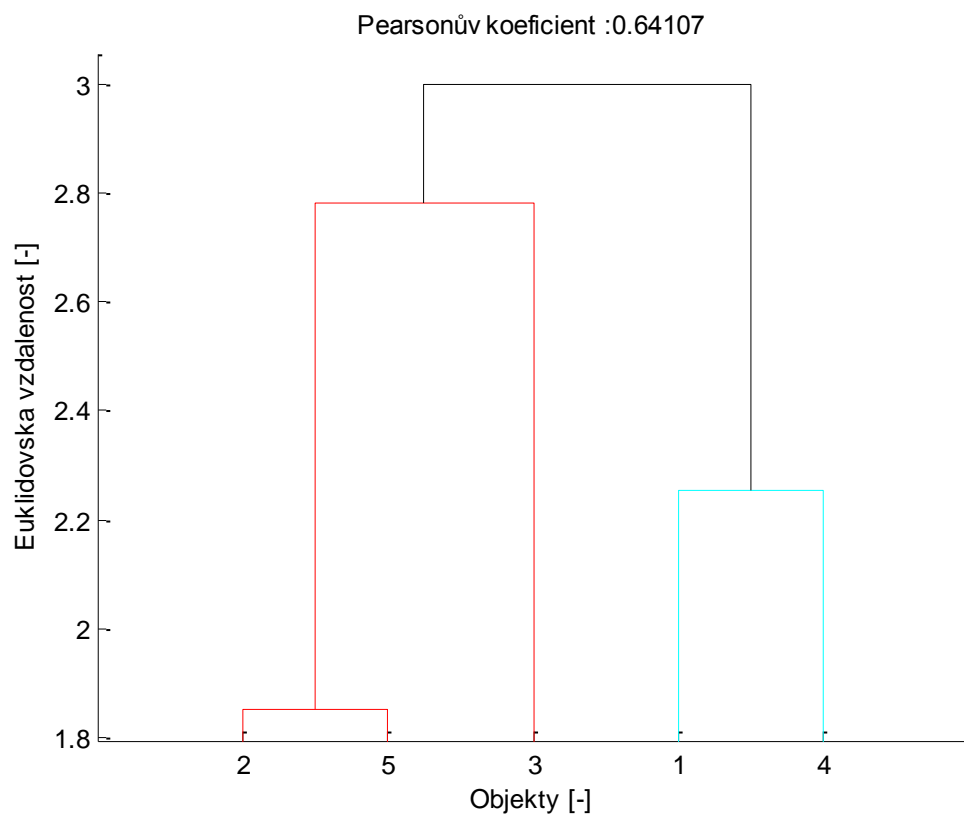
Obrázek 30 - Neúspěšná shluková analýza signálu w006.mat (kombinace atributu plochy a R-R intervalu)



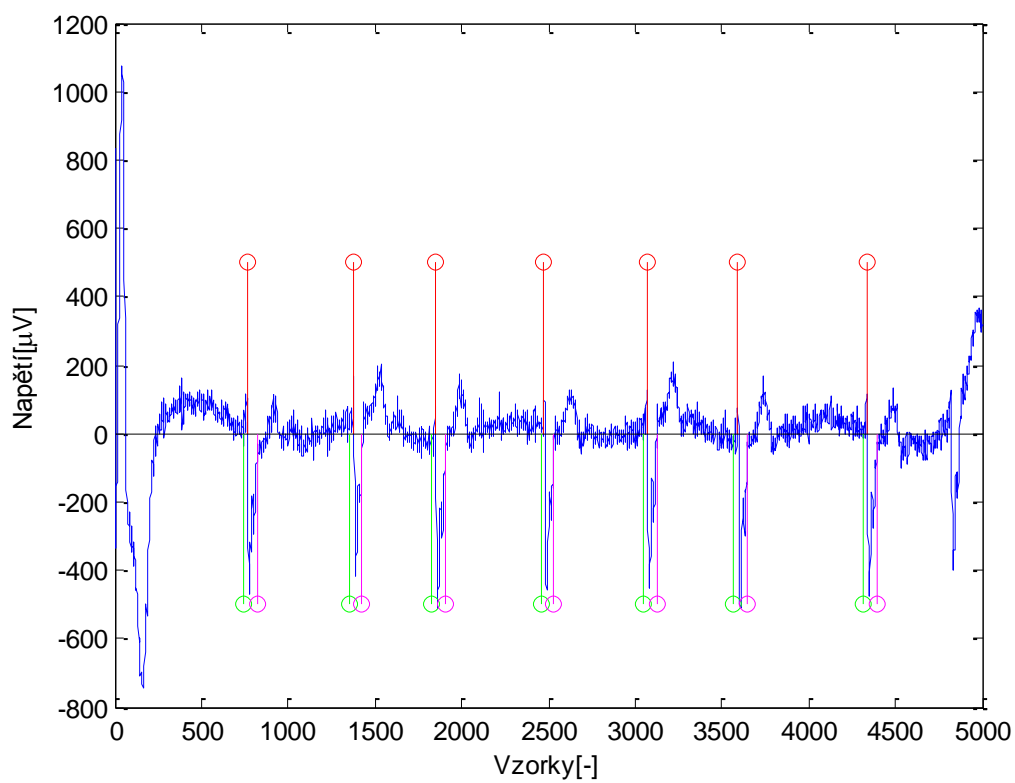
Obrázek 31 - Neúspěšná shluková analýza signálu w006.mat (kombinace atributu tvarového faktoru a R-R intervalu)



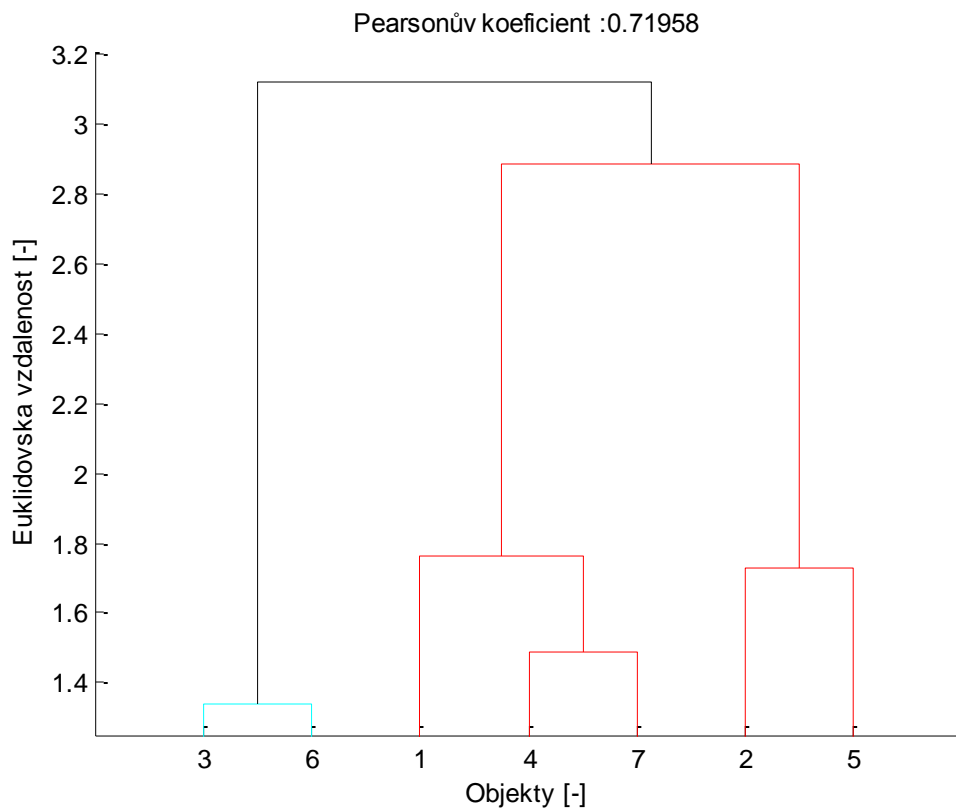
Obrázek 32 - Detekce vln v signálu w054.mat



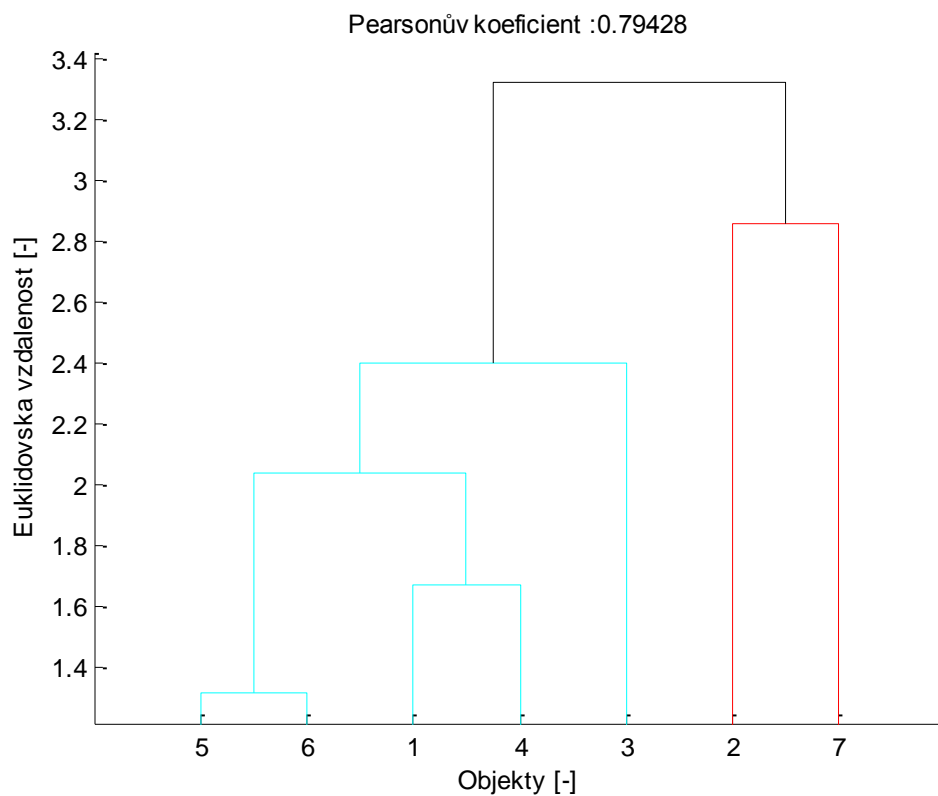
Obrázek 33 - Neúspěšná shluková analýza signálu w054.mat (kombinace atributu plochy a R-R intervalu)



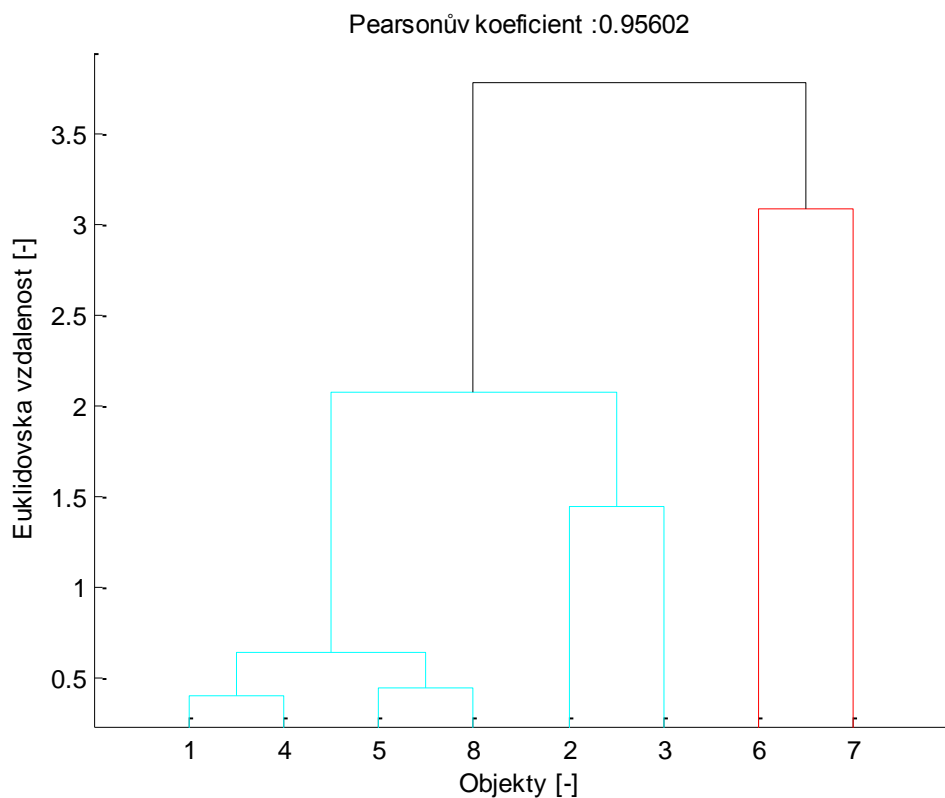
Obrázek 34 - Detekce vln v signálu w120.mat



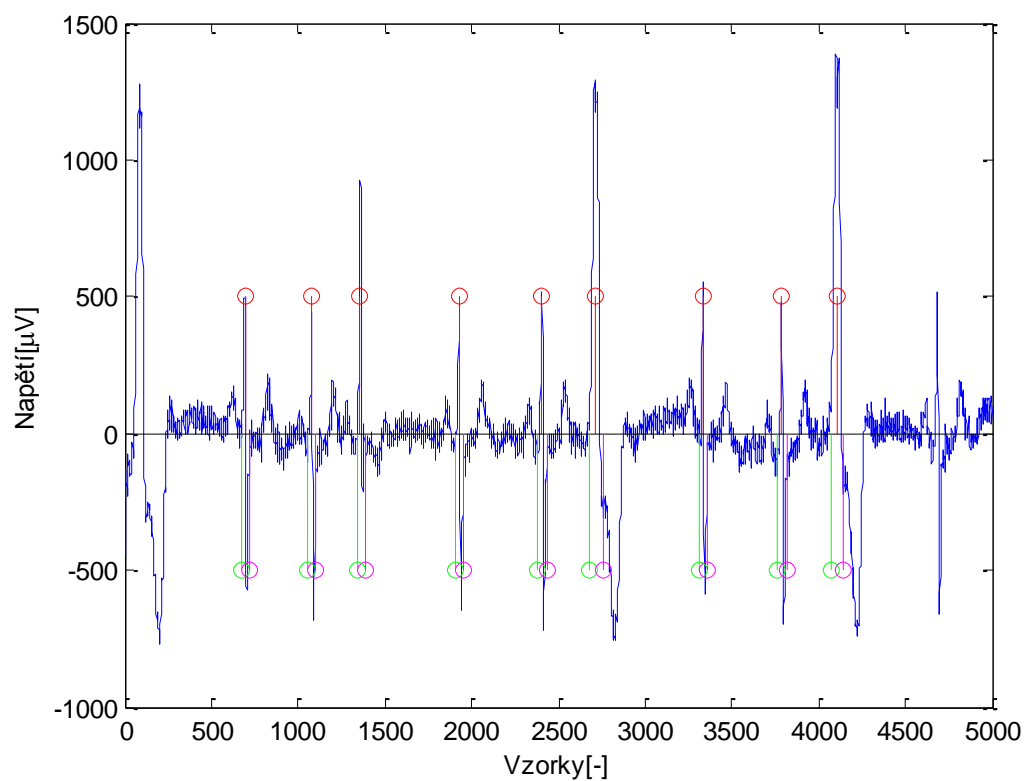
Obrázek 35 - Neúspěšná shluková analýza signálu w120.mat (kombinace atributu plochy a R-R intervalu)



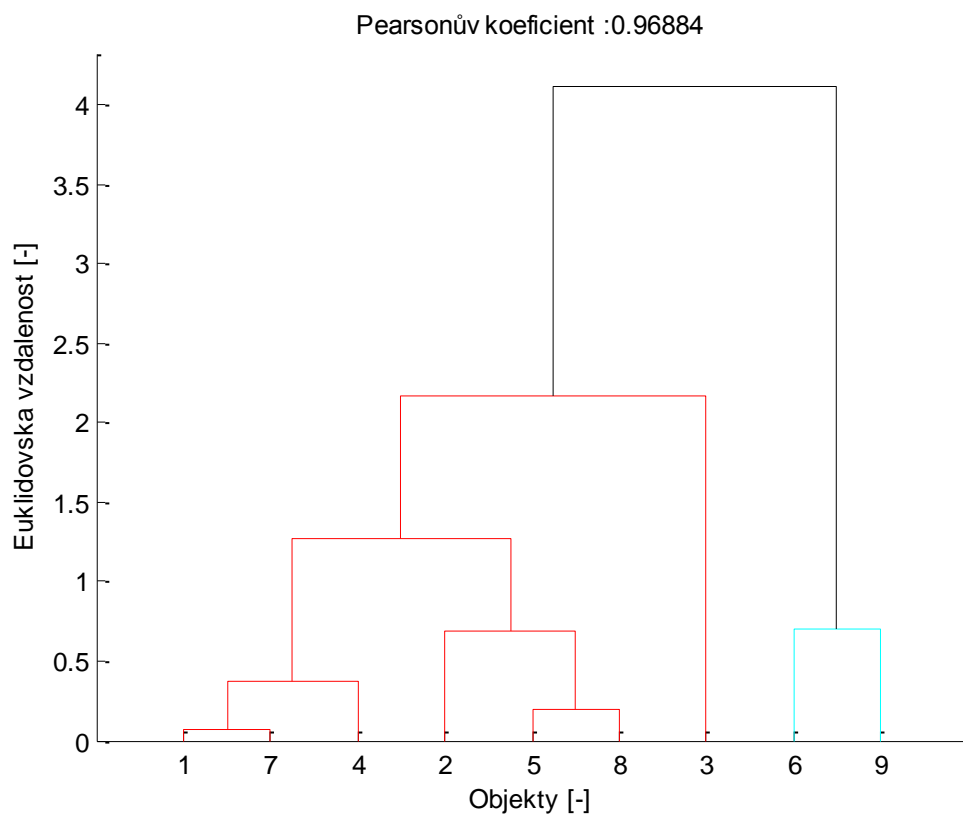
Obrázek 36 - Neúspěšná shluková analýza signálu w120.mat (kombinace atributu tvarového faktoru a R-R intervalu)



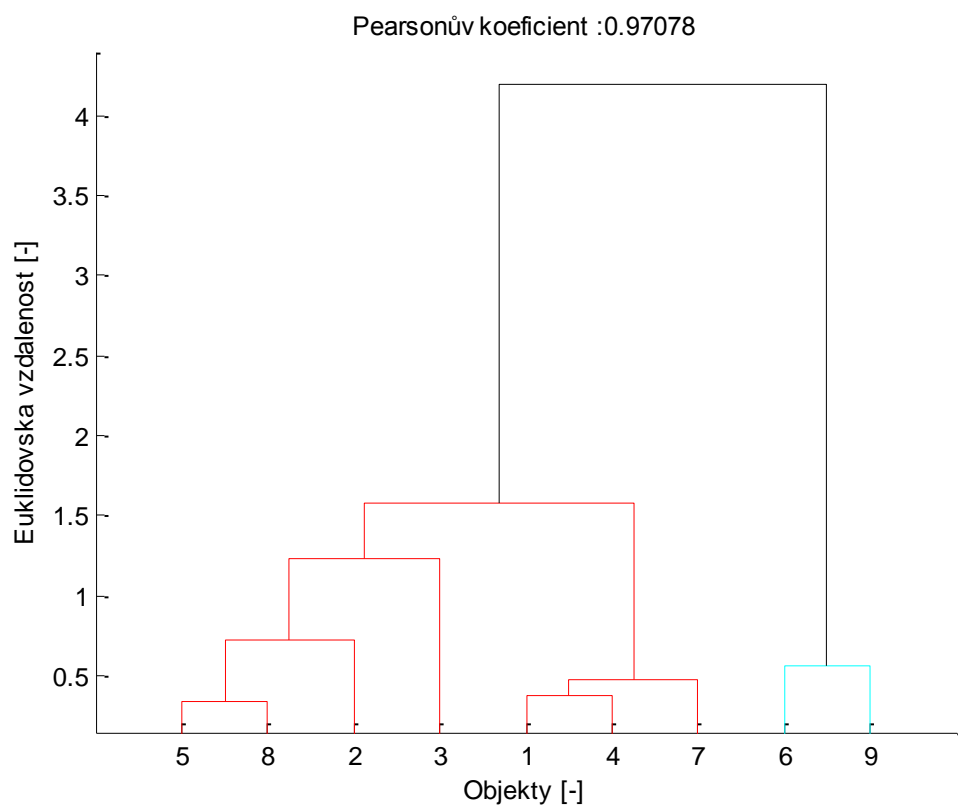
Obrázek 37 - Neúspěšná shluková analýza signálu w028.mat (kombinace atributu tvarového faktoru a R-R intervalu)



Obrázek 38 - Detekce vln v signálu w103.mat

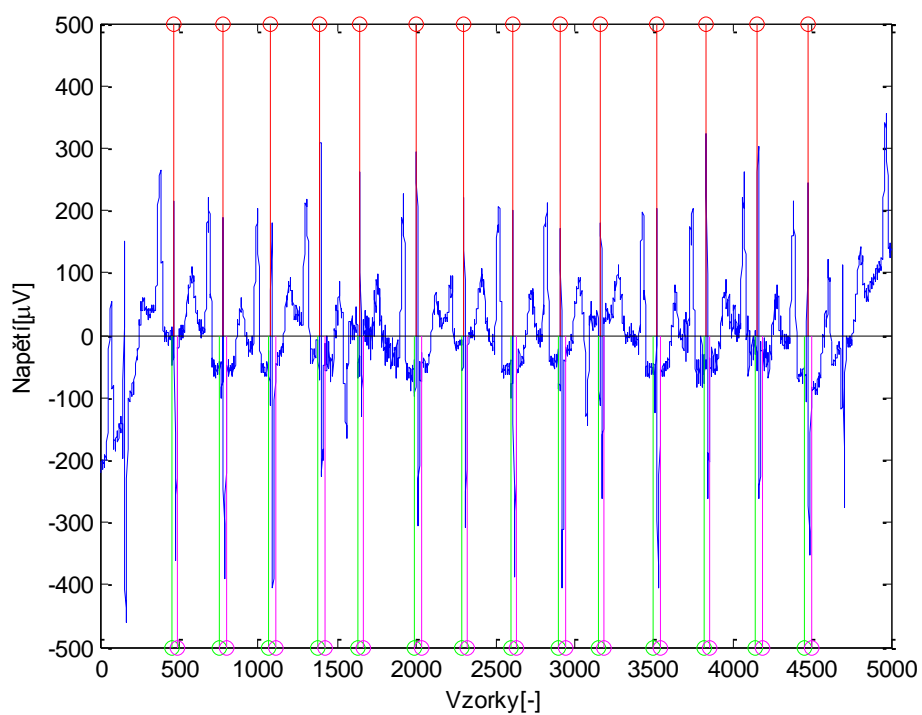


Obrázek 39 - Neúspěšná shluková analýza signálu w103.mat (kombinace atributu plochy a R-R intervalu)

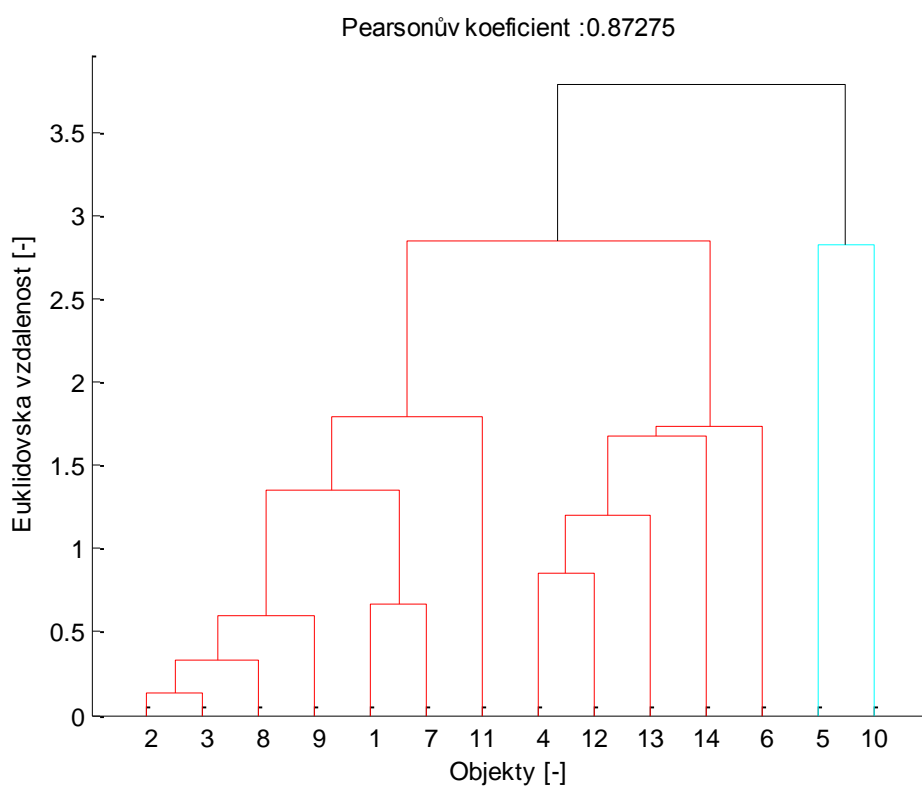


Obrázek 40 - Neúspěšná shluková analýza signálu w103.mat (kombinace atributu tvarového faktoru a R-R intervalu)

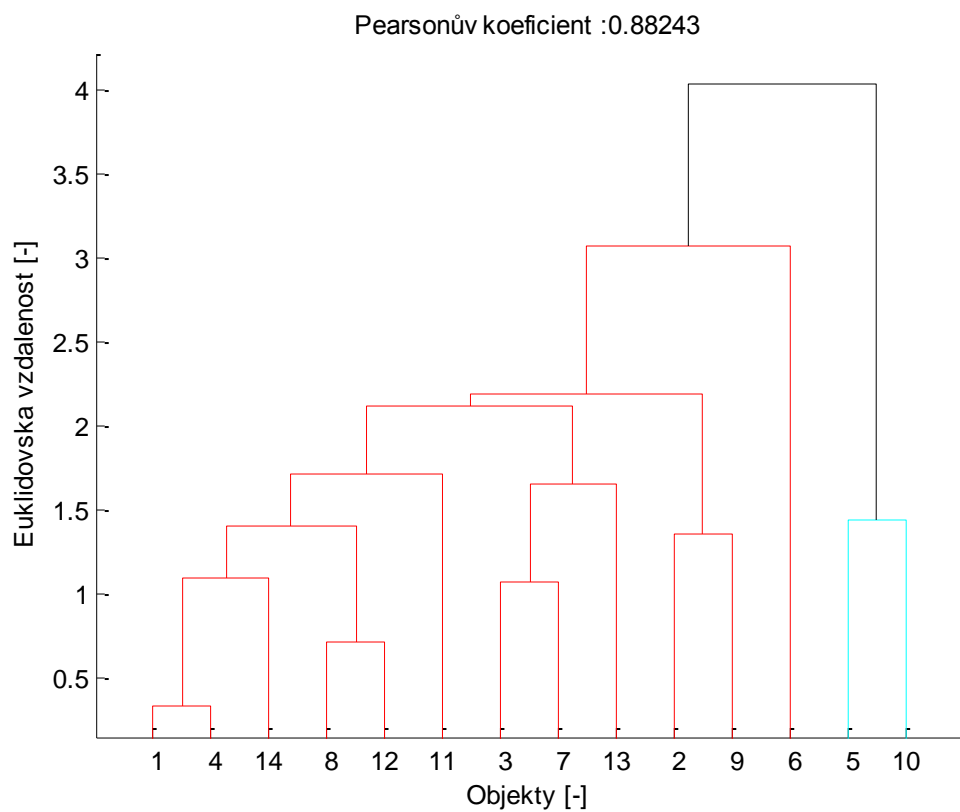
Příloha č. 3 : Další ukázky úspěšné shlukové analýzy pro obě kombinace atributů



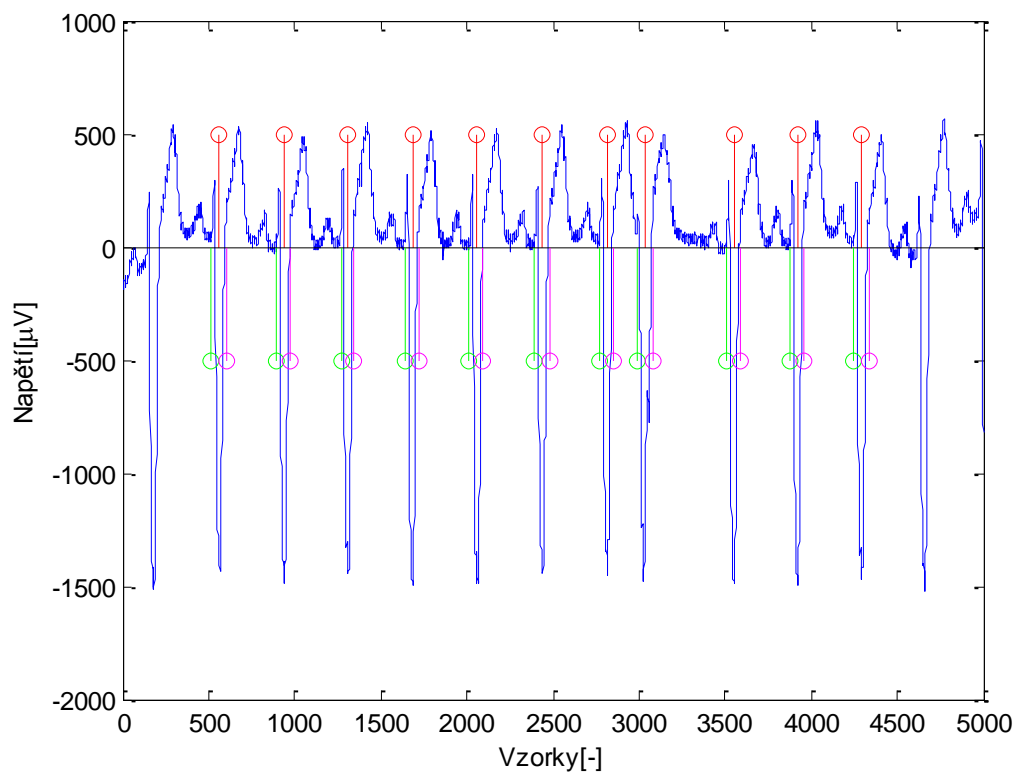
Obrázek 41 - Detekce vln v signálu w047.mat



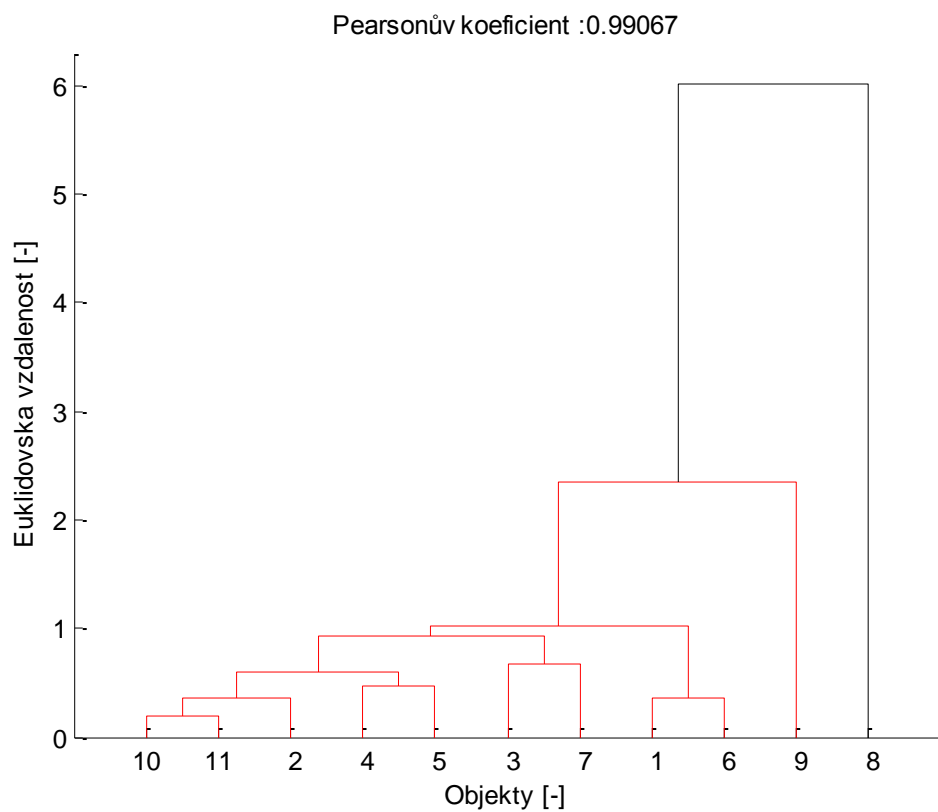
Obrázek 42 - Úspěšná shluková analýza signálu w047.mat (kombinace atributu ploch a R-R intervalu)



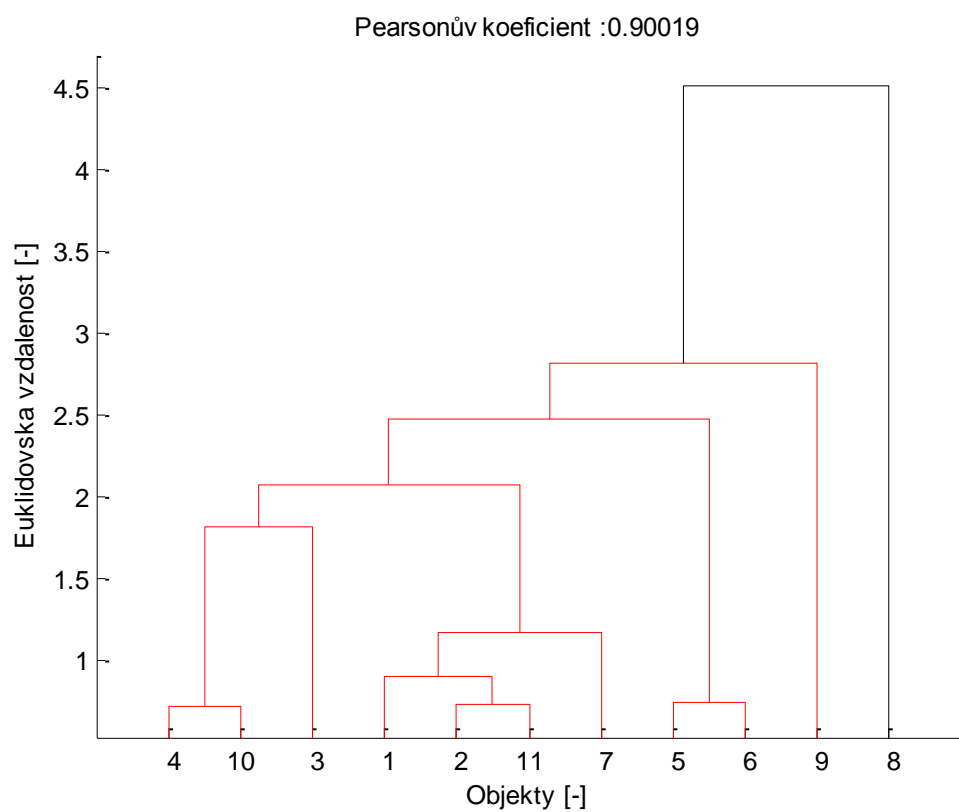
Obrázek 43 - Úspěšná shluková analýza signálu w047.mat (kombinace atributu tvarového faktoru a R-R intervalu)



Obrázek 44 - Detekce vln v signálu w026.mat

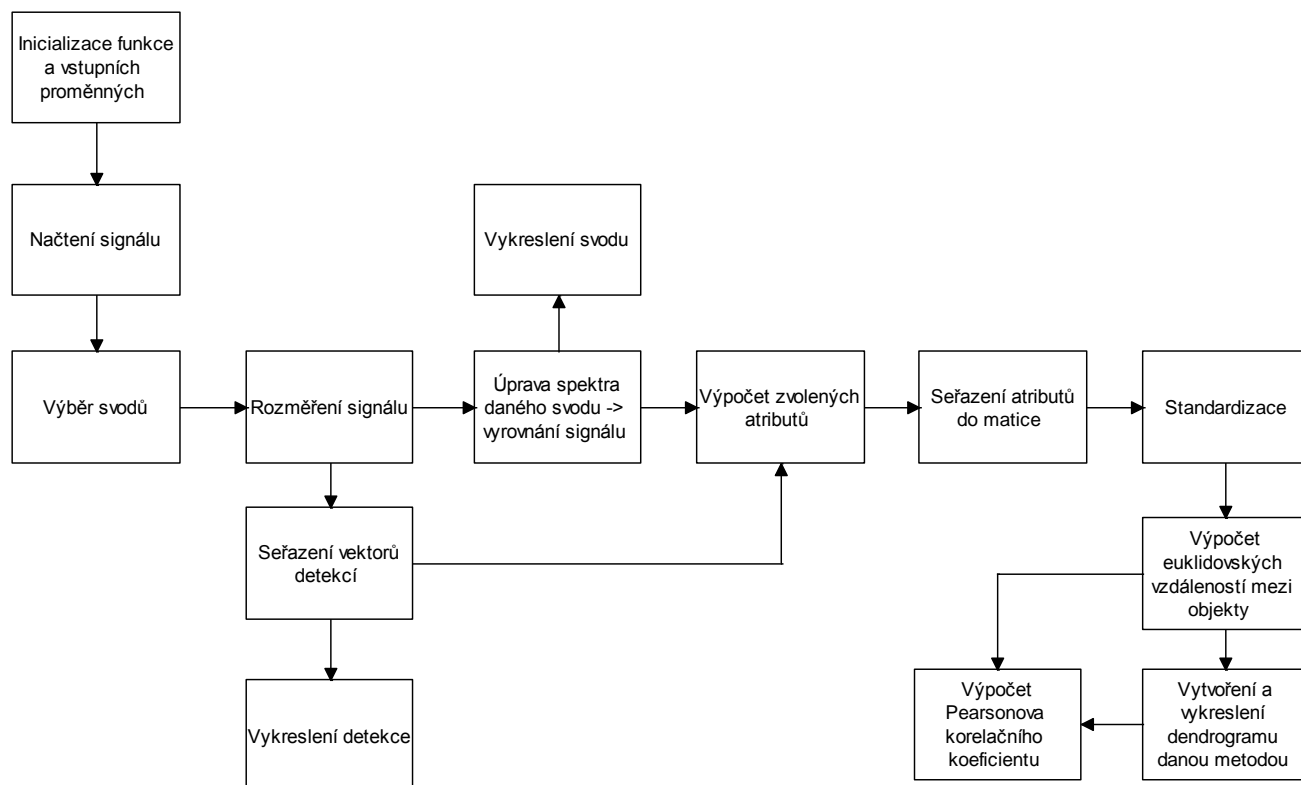


Obrázek 45 - Úspěšná shluková analýza signálu w026.mat (kombinace atributu ploch a R-R intervalu)



Obrázek 46 - Úspěšná shluková analýza signálu w026.mat (kombinace atributu tvarového faktoru a R-R intervalu)

Příloha č. 4 : Blokové schéma programu pro provedení shlukové analýzy



Obrázek 47 - Blokové schéma programu pro provedení shlukové analýzy

Příloha č. 5 : Přílohy na CD

1. Funkce:

shlukovani.m
shlukovani_w117.m
atributy.m
automat.m
FF.m
EKGrozmereni_global.m

2. 21 signálů z knihovny CSE